GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 113.105 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-2
Perfect score: 1238
Sequence: 1 MDKTHTCPPCPAPELLGGPS.....MHEALHNHYTQKSLSLSPGK 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1238	100.0	228	3	AAY96529	Aay96529 Human IgG
2	1238	100.0	228	3	AAB16955	Aab16955 Human IgG
3	1238	100.0	228	4	AAB98953	Aab98953 Human IgG
4	1238	100.0	228	5	ABB04279	Abb04279 Human IgG
5	1238	100.0	228	5	AAU81074	Aau81074 Human IgG
6	1238	100.0	228	5	AAE14310	Aae14310 Human imm
7	1238	100.0	228	5	ABB73410	Abb73410 Human imm
8	1238	100.0	228	5	AAG66012	Aag66012 Human imm
9	1238	100.0	228	5	AAU73018	Aau73018 Human imm
10	1238	100.0	228	6	ABJ38267	Abj38267 Human IgG
11	1238	100.0	228	7	ADN59683	Adn59683 Human IgG
12	1238	100.0	228	8	ADM17708	Adm17708 Human IgG
13	1238	100.0	228	8	ADQ75329	Adq75329 Human IgG
14	1238	100.0	243	3	AAB17957	Aab17957 Fc-MMP in
15	1238	100.0	243	5	ABB73425	Abb73425 Fc-MMP in
16	1238	100.0	243	7	ADN59746	Adn59746 Vector 20
17	1238	100.0	247	3	AAB16958	Aab16958 Fc-TMP pr
18	1238	100.0	247	5	ABB73411	Abb73411 Fc-TPO mi
19	1238	100.0	248	3	AAB17951	Aab17951 Fc-TNF-al
20	1238	100.0	248	3	AAB17953	Aab17953 Fc-IL-1 a
21	1238	100.0	248	5	ABB73421	Abb73421 Fc-interl
22	1238	100.0	248	5	ABB73419	Abb73419 Fc-TNF-al
23	1238	100.0	250	7	ADD31616	Add31616 Ang-2 pep
24	1238	100.0	250	8	ADT71978	Adt71978 Ang-2 neg
25	1238	100.0	252	3	AAB17955	Aab17955 Fc-VEGF a
26	1238	100.0	252	5	ABB73423	Abb73423 Fc-VEGF a
27	1238	100.0	253	3	AAB16964	Aab16964 Fc-EMP pr
28	1238	100.0	253	5	ABB73415	Abb73415 Fc-EPO mi
29	1238	100.0	259	9	AEA18572	Aea18572 Amino aci
30	1238	100.0	268	3	AAB16959	Aab16959 Fc-TMP-TM
31	1238	100.0	268	5	ABB73412	Abb73412 Fc-TMP-TM
32	1238	100.0	269	3	AAY96531	Aay96531 Human IgG
33	1238	100.0	277	3	AAB16967	Aab16967 Fc-EMP-EM
34	1238	100.0	277	5	ABB73418	Abb73418 Fc-EMP-EM
35	1238	100.0	282	5	AAU81169	Aau81169 Echistati
36	1238	100.0	374	2	AAW49075	Aaw49075 Recombina
37	1238	100.0	374	2	AAW83963	Aaw83963 Recombina
38	1238	100.0	374	9	AEB51285	Aeb51285 Recombina
39	1238	100.0	401	4	AAB80904	Aab80904 Human met
40	1238	100.0	401	4	AAY72922	Aay72922 Human met
41	1235	99.8	406	7	ADP75162	Adp75162 Fusion co
42	1235	99.8	409	7	ADP75170	Adp75170 Fusion co
43	1235	99.8	409	7	ADP75176	Adp75176 Fusion co
44	1235	99.8	410	7	ADP75166	Adp75166 Fusion co
45	1235	99.8	412	7	ADP75168	Adp75168 Fusion co

ALIGNMENTS

RESULT 1
AAY96529
ID AAY96529 standard; protein; 228 AA.
XX
AC AAY96529;
XX
DT 04-SEP-2000 (first entry)
XX
DE Human IgG1 Fc chain.
XX
KW Immunoglobulin; IgG1; Fc; thrombopoietin; mimetic; TMP; TPO; platelet;
KW megakaryocyte; production; anti-human immunodeficiency virus; anti-HIV;
KW anti-anaemic; dermatological; immunosuppressive; anti-inflammatory.
XX
OS Homo sapiens.
XX
PN WO200024770-A2.
XX
PD 04-MAY-2000.
XX
PF 22-OCT-1999; 99WO-US024834.
XX
PR 23-OCT-1998; 98US-0105348P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Liu C, Feige U, Cheetham J;
XX
DR WPI; 2000-365108/31.
DR N-PSDB; AAA29220.
XX
PT Thrombopoietic peptides which activate mpl receptors and increase the
PT production of platelets or platelet precursors, useful for treatment of
PT diseases which involve thombocytopenia.
XX
PS Disclosure; Page 76-77; 91pp; English.
XX
CC A compound which binds to an mpl receptor comprising a thrombopoietin
CC mimetic peptide (TMP) dimer joined by a linker [TMP_1-(L_1)_nTMP_2], is
CC new. TMP_1 and TMP_2 are amino acid sequences varying from at least 10 to
CC 14 residues in length comprising X_2-X_1_0, X_2-X_1_1, X_2-X_1_2, X_2-
CC X_1_3, X_2-X_1_4, X_1-X_1_0, X_1-X_1_1, X_1-X_1_2, X_1-X_1_3, and X_1-
CC X_1_4. X_1 = I, A, V, L, S or R; X_2 = E, D, K or V; X_3 = G or A; X_4 =
CC P; X_5 = T or S; X_6 = L, I, V, A or F; X_7 = R or K; X_8 = Q, N, or E;
CC X_9 = W, Y or F; X_1_0 = L, I, V, A, F, M, or K; X_1_1 = A, I, V, L, F,
CC S, T, K, H, or E; X_1_2 = A, I, V, L, F, G, S, or Q; X_1_3 = R, K, T, V,
CC N, Q or G; X_1_4 = A, I, V, L, F, T, R, E, or G; L_1 = linker comprising

CC 1 to 20 amino acids; and n = 0 or 1. The compounds bind to and activate
CC the c-Mpl receptor which mediates the activity of endogenous
CC thrombopoietin. The TMPs are useful for increasing the production of
CC platelets or platelet precursors (e.g. megakaryocytes) in a mammal, which
CC is useful for treatment of diseases which involve thombocytopenia, e.g.
CC aplastic anaemia, immune thrombocytopenia (ITP), human immunodeficiency
CC virus associated ITP, and systemic lupus erythematosus
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 1238; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
```

RESULT 2
AAB16955
ID AAB16955 standard; protein; 228 AA.
XX
AC AAB16955;
XX
DT 31-OCT-2000 (first entry)
XX
DE Human IgG1 Fc protein sequence SEQ ID NO:2.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69443.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 7; Page 176-177; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 1238; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
```

RESULT 3
AAB98953
ID AAB98953 standard; protein; 228 AA.
XX
AC AAB98953;
XX
DT 14-AUG-2001 (first entry)
XX
DE Human IgG1 Fc region.
XX
KW Human; IgG1; immunoglobulin; Fc region; Fc fusion protein; misfolding;
KW therapy; cancer; osteoarthritis; AIDS; obesity; inflammation;
KW transplant rejection.
XX
OS Homo sapiens.
XX
PN WO200134638-A1.
XX
PD 17-MAY-2001.
XX
PF 10-NOV-2000; 2000WO-US030798.
XX
PR 12-NOV-1999; 99US-0165188P.
PR 09-NOV-2000; 2000US-00709704.
XX
PA (AMGE-) AMGEN INC.
XX
PI Treuheit MJ, O'conner SR, Kosky AA;
XX
DR WPI; 2001-335908/35.
DR N-PSDB; AAH25762.
XX

PT Correcting disulfide bond misfolds in Fc-containing proteins,
PT particularly therapeutic Fc-containing fusion proteins or antibodies, by
PT treatment with copper halide.
XX
PS Claim 30; Fig 5; 59pp; English.
XX
CC The present invention describes a process for preparing a
CC pharmacologically active compound, involving preparing a compound
CC comprising an immunoglobulin Fc domain fused to a protein of interest,
CC treating the compound with a copper(II) halide and isolating the treated
CC molecule. This can be used to correct misfolding of Fc domain containing
CC proteins, for use in therapeutic agents which may be used in the
CC treatment of cancer, inflammation, transplant rejection, AIDS,
CC osteoarthritis and obesity. The present sequence is the IgG1 Fc domain
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 1238; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
```

RESULT 4
ABB04279
ID ABB04279 standard; protein; 228 AA.
XX
AC ABB04279;
XX
DT 13-FEB-2002 (first entry)
XX
DE Human IgG1 Fc domain.
XX
KW Glucagon antagonist; antidiabetic; anti-hormonal; Fc domain;
KW non-insulin dependent diabetes mellitus; human; immunoglobulin G; IgG.
XX
OS Homo sapiens.
XX
PN WO200183527-A2.
XX
PD 08-NOV-2001.
XX
PF 03-MAY-2001; 2001WO-US014321.
XX
PR 03-MAY-2000; 2000US-0201436P.
PR 02-MAY-2001; 2001US-00847249.
XX
PA (AMGE-) AMGEN INC.
XX
PI Marshall WS, Stark KL;
XX
DR WPI; 2002-017738/02.
DR N-PSDB; ABA03672.
XX
PT Compositions comprising glucagon antagonist domains, useful for treating
PT diabetes mellitus.
XX
PS Claim 8; Fig 2; 54pp; English.
XX
CC The invention relates to compositions comprising a glucagon antagonist
CC domain and a vehicle, such as a polymer (e.g. PEG or dextran) or,
CC preferably, an Fc domain. The vehicle is covalently attached to the
CC glucagon antagonist domain. The compositions are administered to treat
CC non-insulin dependent diabetes mellitus. The present sequence is the
CC human IgG Fc domain, which may be used as the vehicle in the compositions
CC of the invention
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 1238; DB 5; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
```

RESULT 5
AAU81074
ID AAU81074 standard; protein; 228 AA.
XX
AC AAU81074;
XX
DT 09-APR-2002 (first entry)
XX
DE Human IgG1 Fc.
XX
KW Human; IgG Fc; anticoagulant; thrombolytic; cytostatic; antiinflammatory;
KW immunosuppressive; osteopathic; antagonist; laminin; saw-scaled viper;
KW echistatin; integrin; selectin; vinculin; platelet aggregation;
KW angiogenesis; tumour; inflammation; autoimmune disease;
KW rheumatoid arthritis; osteoporosis.
XX
OS Homo sapiens.
XX
PN WO200181377-A2.
XX
PD 01-NOV-2001.
XX
PF 23-APR-2001; 2001WO-US013069.
XX
PR 21-APR-2000; 2000US-0198919P.
PR 03-MAY-2000; 2000US-0201394P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Kohno T, Lacey DL, Boone TC;
XX
DR WPI; 2002-062025/08.
DR N-PSDB; ABK24097.
XX
PT Composition comprising integrin or adhesion antagonistic peptide and
PT vehicle, useful for treating or preventing platelet aggregation, has a
PT longer half-life than free peptide.
XX
PS Claim 9; Fig 3; 68pp; English.
XX
CC The invention relates to a composition comprising an integrin/adhesion

```
CC   antagonistic peptide (I) and a vehicle e.g. IgG Fc. The peptides are
CC   based on laminin or saw-scaled viper echistatin and target integrin,
CC   selectin or vinculin.Also included are compounds of formula (Ia) and
CC   their multimers (X^1)_a-F^1-(X^2)_b where; F^1 = Fc domain; X^1 and X^2 =
CC   -(L^1)_c-P^1, (L^1)_c-P^1-(L^2)_d-P^2, (L^1)_c-P^1-(L^2)_d-P^2-(L^3)_e-
CC   P^3 or (L^1)_c-P^1-(L^2)_d-P^2-(L^3)_e-P^3-(L^4)_f-P^4; P^1-P^4 = same or
CC   different (I); L^1-L^4 = same or different linkers; a-f = 0 or 1,
CC   provided at least one of a and b = 1, a nucleic acid that encodes (Ia),
CC   an expression vector containing the nucleic acid, host cells containing
CC   the vector, producing a pharmaceutically active compound (B) by
CC   covalently linking at least one Fc domain to at least one amino acid
CC   sequence of a selected randomized (I) and any of six laminin-related
CC   peptides (Ib). The compositions are used prophylactically and
CC   therapeutically in the same way as (I), e.g. to inhibit platelet
CC   aggregation or angiogenesis (tumours), or to treat inflammation and
CC   autoimmune diseases (e.g. rheumatoid arthritis) and many different forms
CC   of osteoporosis, also for diagnosis. Attaching the vehicle (especially Fc
CC   domain) to (I) increases the half-life (free (I) are normally degraded
CC   very quickly in vivo). The present sequence is human IgG1 Fc which is
CC   used as a vehicle for the antagonists of the invention
XX
SQ   Sequence 228 AA;

  Query Match             100.0%;  Score 1238;  DB 5;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-90;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228


RESULT 6
AAE14310
ID   AAE14310 standard; protein; 228 AA.
XX
AC   AAE14310;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Human immunoglobulin G (IgG1) Fc.
XX
KW   Human; calcitonin; CT; CT receptor; Fc domain; therapy; osteoporosis;
KW   immunoglobulin G; IgG; osteopathic.
XX
OS   Homo sapiens.
XX
PN   WO200183526-A2.
XX
PD   08-NOV-2001.
XX
PF   03-MAY-2001; 2001WO-US014320.
XX
PR   03-MAY-2000; 2000US-0201511P.
PR   02-MAY-2001; 2001US-00847712.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Liu C,  Marshall WS,  Reynolds A;
XX
DR   WPI; 2002-034503/04.
DR   N-PSDB; AAD23840.
XX
PT   Compositions comprising Calcitonin receptor modulator domains, useful for
PT   treating osteoporosis.
XX
PS   Claim 8; Fig 3; 64pp; English.
XX
CC   The invention relates to therapeutic agents that modulate the activity of
CC   calcitonin (CT) receptor. Modulators of CT receptor comprise a CT
CC   receptor modulating domain and a vehicle such as a polymer or an Fc
CC   domain, where the vehicle is covalently attached to the CT receptor
CC   modulating domain. The compositions comprising CT receptor modulating
CC   domains are used to treat osteoporosis. The present sequence is human
CC   immunoglobulin G (IgG1) Fc protein used in the invention
XX
SQ   Sequence 228 AA;

  Query Match             100.0%;  Score 1238;  DB 5;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-90;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228


RESULT 7
ABB73410
ID   ABB73410 standard; protein; 228 AA.
XX
AC   ABB73410;
XX
DT   05-APR-2002  (first entry)
XX
DE   Human immunoglobulin G1 Fc (IgG1 Fc) amino acid SEQ ID NO:2.
XX
KW   Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW   erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW   TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW   TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW   MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW   cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW   antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW   neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW   cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW   sleep disorder; neurological degenerative disease; anaemia;
KW   thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW   Fanconi's syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200183525-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014310.
XX
PR   03-MAY-2000; 2000US-00563286.
XX
PA   (AMGE-) AMGEN INC.
XX
```

PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35760.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 7; Fig 4; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 228 AA;

```
  Query Match             100.0%;  Score 1238;  DB 5;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-90;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
```

RESULT 8
AAG66012
ID AAG66012 standard; protein; 228 AA.
XX
AC AAG66012;
XX
DT 27-FEB-2002 (first entry)
XX
DE Human immunoglobulin (Ig) G1 Fc region sequence.
XX
KW Apo-AI; amphipathic; pharmaceutical; peptide mimic; antilipemic;
KW anti-HIV; virucide; immunoglobulin; IgG1.
XX
OS Homo sapiens.
XX
PN WO200181376-A2.
XX
PD 01-NOV-2001.
XX
PF 23-APR-2001; 2001WO-US013068.
XX
PR 21-APR-2000; 2000US-0198920P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Kohno T;
XX
DR WPI; 2002-049262/06.
DR N-PSDB; AAI67658.
XX
PT Recombinant or modified therapeutic agents having Apo-AI amphipathic
PT helix peptide activity useful in treatment of hypercholesterolemia and
PT viral infections such as herpes simplex virus, human immunodeficiency
PT virus.
XX
PS Claim 8; Fig 3A-B; 49pp; English.
XX
CC The invention provides a composition comprising a therapeutic agent that
CC has activity similar to Apo-AI amphipathic helix peptide, but with better
CC pharmaceutical characteristics attached to a vehicle through the
CC peptide's N-terminus or C-terminus having a specified formula. The
CC peptide mimic has greater half-life compared to conventional Apo-AI
CC amphipathic helix peptide. The compositions are useful for treating
CC hypercholesterolemia and viral infection such as HIV, HSV. The present
CC sequence represents the human immunoglobulin (Ig) G1 Fc region which acts
CC as a vehicle
XX
SQ Sequence 228 AA;

```
  Query Match             100.0%;  Score 1238;  DB 5;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-90;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
```

RESULT 9
AAU73018
ID AAU73018 standard; protein; 228 AA.
XX
AC AAU73018;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human immunoglobulin G (IgG) Fc region.
XX
KW Human; parathyroid hormone; PTH; parathyroid hormone-related protein;
KW PTHrP; bone resorption inhibitor; osteoprotegrin; OPG; OPG-L antibody;
KW calcitonin; bisphosphonate; oestrogen; oestrogen receptor; tibolone;
KW osteopenia; hyperthyroidism; hypercalcaemia; tumour metastasis; bone;
KW breast cancer; prostate cancer; cachexia; anorexia; osteoporosis;
KW Paget's disease; osteomyelitis; osteonecrosis; bone cell death;
KW Gaucher's disease; sickle cell anaemia; systemic lupus erythematosus;
KW rheumatoid arthritis; periodontal disease; alopecia; fracture repair;
KW immunoglobulin G; IgG.

XX
OS Homo sapiens.
XX
PN WO200181415-A2.
XX
PD 01-NOV-2001.
XX
PF 27-APR-2001; 2001WO-US013528.
XX
PR 27-APR-2000; 2000US-0200053P.
PR 28-JUN-2000; 2000US-0214860P.
PR 06-FEB-2001; 2001US-0266673P.
PR 26-APR-2001; 2001US-00843221.
XX
PA (AMGE-) AMGEN INC.
XX
PI Kostenuik P, Liu C, Lacey DL;
XX
DR WPI; 2002-066435/09.
DR N-PSDB; AAS97392.
XX
PT Composition, useful for treating osteopenia, comprises parathyroid
PT hormone and parathyroid hormone-related protein receptor modulators.
XX
PS Claim 6; Fig 3; 107pp; English.
XX
CC The invention relates to a composition (I) comprising modulators of
CC parathyroid hormone (PTH) and parathyroid hormone-related protein (PTHrP)
CC which comprise a PTH/PTHrP modulating domain and a vehicle. (I)
CC comprising PTH agonist optionally with a bone resorption inhibitor, such
CC as osteoprotegrin (OPG), OPG-L antibody, calcitonin, bisphosphonates,
CC oestrogens, oestrogen receptor modulators and tibolone is useful for
CC treating osteopenia. (I) is useful for therapeutic and prophylactic
CC purposes. Antagonists of PTH receptor are useful in treating primary and
CC secondary hyperthyroidism, hypercalcaemia, tumour metastases,
CC particularly breast and prostate cancer, cachexia and anorexia,
CC osteopenia, including various forms of osteoporosis, Paget's disease of
CC bone, osteomyelitis, osteonecrosis or bone cell death, associated with
CC traumatic injury or nontraumatic necrosis associated with Gaucher's
CC disease, sickle cell anaemia, systemic lupus erythematosus, rheumatoid
CC arthritis, periodontal disease and alopecia. PTH receptor agonists are
CC useful as therapeutic agents in conditions including fracture repair
CC (including healing of non-union fractures), osteopenia, including various
CC forms of osteoporosis. AAU73018-AAU73181 represent parathyroid hormone
CC and parathyroid hormone related protein (PTH/PTHrP) modulators and
CC related amino acid sequences of the invention
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 1238; DB 5; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
Db 1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
Db 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
Db 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy 181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
Db 181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228

RESULT 10
ABJ38267
ID ABJ38267 standard; protein; 228 AA.

XX
AC ABJ38267;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human IgG1 Fc protein SEQ ID No 2.
XX
KW TALL-1-binding protein; TALL-1; B-cell-mediated autoimmune disease;
KW systemic lupus erythematosus; B-cell-mediated cancer; lymphoma;
KW inflammation; rheumatoid arthritis; acute pancreatitis; atherosclerosis;
KW Alzheimer's disease; asthma; cachexia; cirrhosis; diabetes; osteoporosis;
KW glomerulonephritis; Hashimoto's thyroiditis; ischaemic injury; psoriasis;
KW multiple myeloma; multiple sclerosis; Parkinson's disease; vasculitis;
KW gene therapy; human IgG1Fc; human.
XX
OS Homo sapiens.
XX
PN WO200292620-A2.
XX
PD 21-NOV-2002.
XX
PF 13-MAY-2002; 2002WO-US015273.
XX
PR 11-MAY-2001; 2001US-0290196P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Min H, Hsu H;
XX
DR WPI; 2003-156719/15.
DR N-PSDB; ABT33856.
XX
PT New TALL-1-binding polypeptide, useful for modulating the activity of
PT TALL-1 and in treating, preventing or diagnosing a B-cell-mediated
PT autoimmune diseases, cancers or lymphomas.
XX
PS Claim 36; Fig 3; 236pp; English.
XX
CC The invention relates to a novel TALL-1-binding polypeptide comprising a
CC defined sequence in the specification. The composition is useful in
CC modulating the activity of TALL-1, and in treating, preventing,
CC ameliorating, diagnosing or prognosing a B-cell-mediated autoimmune
CC disease (e.g. systemic lupus erythematosus) or B-cell-mediated cancer or
CC lymphoma. The composition may also be used in treating inflammations
CC (e.g. rheumatoid arthritis), acute pancreatitis, Alzheimer's disease,
CC asthma, atherosclerosis, cachexia, cirrhosis, diabetes,
CC glomerulonephritis, Hashimoto's thyroiditis, ischaemic injury, multiple
CC myeloma, multiple sclerosis, osteoporosis, Parkinson's disease, psoriasis
CC and vasculitis. Disorders may be treated with the novel composition using
CC gene therapy. This sequence represents a human IgG1Fc protein relating to
CC the TALL-1 sequence of the invention
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 1238; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
Db 1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
Db 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
Db 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy 181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
Db 181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228

RESULT 11
ADN59683
ID ADN59683 standard; protein; 228 AA.
XX
AC ADN59683;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human IgG1 Fc amino acid sequence, seq id 32.
XX
KW Haemostatic; antianaemic; immunosuppressive; platelet;
KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;
KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;
KW thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;
KW autoimmune haemolytic anaemia; Hughe's syndrome;
KW lupoid thrombocytopaenia; IgG1.
XX
OS Homo sapiens.
XX
PN WO2003031589-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032552.
XX
PR 11-OCT-2001; 2001US-0328666P.
PR 10-OCT-2002; 2002US-00269806.
XX
PA (AMGE-) AMGEN INC.
XX
PI Min H, Sitney KC, Hartley C;
XX
DR WPI; 2003-403101/38.
DR N-PSDB; ADN59682.
XX
PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and
PT which stimulate the production of platelets and/or the production of
PT platelet precursors, useful for treating thrombocytopenia.
XX
PS Disclosure; SEQ ID NO 32; 126pp; English.
XX
CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that
CC binds to the c-mpl (mpl) receptor, and which stimulates the production of
CC platelets and/or the production of platelet precursors, is new. Further
CC disclosed is a composition of matter (II) that binds to an mpl receptor,
CC and a pharmaceutical composition comprising (II) and a carrier. The
CC pharmaceutical composition of the invention is useful for treating
CC thrombocytopaenia in an animal, and for increasing megakaryocytes or
CC platelets in a patient. The TMP of the invention is useful for treating
CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.
CC disease conditions involving thrombocytopaenia such as aplastic anaemia,
CC autoimmune thrombocytopaenia, drug induced immune thrombocytopaenia,
CC autoimmune haemolytic anaemia, Hughe's syndrome and lupoid
CC thrombocytopaenia. The TMP of the invention is also useful for
CC maintaining the viability or storage life of platelets and/or
CC megakaryocytes and its derived cells. The compounds demonstrate an
CC improved ability to bind to and/or trigger transmembrane signal through,
CC i.e. activating, the mpl receptor the compounds have superior
CC thromobopoietic activity, i.e. the ability to stimulate, in vivo and in
CC vitro, the production of platelets and/or megakaryocytopoietic activity,
CC i.e. the ability to stimulate, in vivo and in vitro, the production of
CC platelet precursors. Further, certain of the compounds also exhibit
CC superior therapeutic properties, such as improved plasma half-life,
CC biological activity and in vivo circulation time. The current sequence
CC represents the human IgG1 Fc protein that may be used as a preferred
CC vehicle of the invention.
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 1238; DB 7; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.6e-90;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
```

RESULT 12
ADM17708
ID ADM17708 standard; protein; 228 AA.
XX
AC ADM17708;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human IgG1 Fc protein SEQ ID NO:60.
XX
KW nerve growth factor modulator; NGF modulator; analgesic; NGF inhibitor;
KW nerve growth factor inhibitor; neurologic pain; diabetic neuropathy;
KW post-herpetic neuralgia; inflammatory pain; migraine; asthma;
KW hyperactive bladder; psoriasis; cancer; acute pain; dental pain;
KW surgical pain; pain; causalgia; demyelinating disease;
KW trigeminal neuralgia; chronic alcoholism; stroke; thalamic pain syndrome;
KW diabetes; acquired immuno deficiency syndrome; AIDS; headache;
KW inflammation; arthritis; rheumatic disease; lupus; osteoarthritis;
KW inflammatory bowel disorder; inflammatory eye disorder; sunburn;
KW carditis; dermatitis; myositis; neuritis; collagen vascular disease;
KW chronic inflammatory condition; neuropathic pain; genitourinary; wound;
KW burn; allergic skin reaction; pruritus; vitiligo;
KW gastrointestinal disorder; colitis; gastric ulceration; duodenal ulcer;
KW human; IgG1 Fc; immunoglobulin G.
XX
OS Homo sapiens.
XX
PN WO2004026329-A1.
XX
PD 01-APR-2004.
XX
PF 19-SEP-2003; 2003WO-US029866.
XX
PR 19-SEP-2002; 2002US-0412524P.
PR 18-SEP-2003; 2003US-00666480.
XX
PA (AMGE-) AMGEN INC.
XX
PI Boone TC, Wild KD, Sitney KC, Min H, Kimmel B;
XX
DR WPI; 2004-283150/26.
DR N-PSDB; ADM17707.
XX
PT Novel peptide capable of modulating nerve growth factor activity, useful
PT for treating disease or disorder e.g., acute pain, dental pain, cancer,
PT migraine and collagen vascular disease.
XX
PS Claim 16; SEQ ID NO 60; 267pp; English.
XX
CC The present invention describes a peptide (I) that is capable of
CC modulating nerve growth factor (NGF) activity. Also described: (1)
CC modified peptide (II) comprising (I) and a vehicle, where the modified
CC peptide is capable of modulating NGF activity; (2) dimer or multimer of
CC (I); (3) modified peptide (III), its multimers or its salt, where the

```
CC   peptide is capable of modulating NGF activity; (4) polynucleotide (IV)
CC   encoding (I), (II) or (III); (5) expression vector (V) comprising (IV);
CC   (6) host cell (VI) comprising (V); (7) a composition (VII) of matter and
CC   a vehicle, where the composition of matter is capable of modulating NGF
CC   activity; and (8) pharmaceutical composition comprising (I), (II) or
CC   (III) and a diluent or carrier. (I) has analgesic activity, and can be
CC   used as an inhibitor of NGF. (I) is useful for treating or preventing a
CC   disease or disorder associated with NGF activity by administering (I) to
CC   human or animal. The disease or disorder chosen from neurologic pain,
CC   painful diabetic neuropathy, post-herpetic neuralgia, inflammatory pain,
CC   migraine, asthma, hyperactive bladder, psoriasis, cancer, acute pain,
CC   dental pain, pain from trauma, surgical pain, pain resulting from
CC   amputation or abscess, causalgia, demyelinating diseases, trigeminal
CC   neuralgia, chronic alcoholism, stroke, thalamic pain syndrome, diabetes,
CC   acquired immuno deficiency syndrome (AIDS), toxins and chemotherapy,
CC   general headache, cluster headache, mixed-vascular and non-vascular
CC   syndromes, tension headache, general inflammation, arthritis, rheumatic
CC   diseases, lupus, osteoarthritis, inflammatory bowel disorders,
CC   inflammatory eye disorders, inflammatory or unstable bladder disorders,
CC   skin complaints with inflammatory components, sunburn, carditis,
CC   dermatitis, myositis, neuritis, collagen vascular diseases, chronic
CC   inflammatory conditions, inflammatory pain associated hyperalgesia and
CC   allodynia, neuropathic pain and associated hyperalgesia and allodynia,
CC   diabetic neuropathy pain, sympathetically maintained pain,
CC   deafferentation syndromes, epithelial tissue damage or dysfunction,
CC   herpes simplex, post-herpetic neuralgia, disturbances of visceral
CC   motility at respiratory, genitourinary, gastrointestinal or vascular
CC   regions, wounds, burns, allergic skin reactions, pruritus, vitiligo,
CC   general gastrointestinal disorders, colitis, gastric ulceration, duodenal
CC   ulcers, vasomotor or allergic rhinitis, or bronchial disorders. (I) is
CC   also useful for modulating pain or promoting analgesia by administering
CC   (I) to human or animal. (I) is also useful in the manufacture of
CC   medicament for the treatment of disease or disorder. The present sequence
CC   is used in the exemplification of the present invention.
XX
SQ   Sequence 228 AA;

  Query Match             100.0%;  Score 1238;  DB 8;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-90;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228


RESULT 13
ADQ75329
ID   ADQ75329 standard; protein; 228 AA.
XX
AC   ADQ75329;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Human IgG1 Fc protein.
XX
KW   parathyroid hormone; parathyroid hormone-related protein; PTH; PTHrP;
KW   osteopathic; osteopenia; IgG Fc; antibody.
XX
OS   Homo sapiens.
XX
PN   WO2004060386-A1.
XX
PD   22-JUL-2004.
XX
PF   01-NOV-2002; 2002WO-US036419.
XX
PR   01-NOV-2002; 2002WO-US036419.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Kostenuik P,  Gegg CV,  Jarosinski MA,  Kinstler OB;
XX
DR   WPI; 2004-543796/52.
XX
PT   New composition of matter comprising parathyroid hormone/parathyroid
PT   hormone-related protein (PTH/PTHrP) modulating domain and a vehicle, or
PT   its multimers, useful for treating osteopenia.
XX
PS   Disclosure; Fig 3A-C; 132pp; English.
XX
CC   The invention relates to a composition comprising the formula (I): (I) P1
CC   -(L1)a-F1, where F1 = a vehicle and is attached at the C-terminus of P1-
CC   (L1)a or through a sidechain at any residue from residue 14 through the C
CC   -terminal residue; P1 = a parathyroid hormone/parathyroid hormone-related
CC   protein (PTH/PTHrP) modulating domain; L1 is a linker; and a = 0 or 1.
CC   The composition of matter is useful for treating osteopenia. This
CC   sequence corresponds to a human IgG Fc used in the invention.
XX
SQ   Sequence 228 AA;

  Query Match             100.0%;  Score 1238;  DB 8;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-90;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228


RESULT 14
AAB17957
ID   AAB17957 standard; protein; 243 AA.
XX
AC   AAB17957;
XX
DT   31-OCT-2000  (first entry)
XX
DE   Fc-MMP inhibitor fusion protein sequence SEQ ID NO:1068.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW   inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW   thrombosis; pharmaceutical.
XX
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
```

PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69507.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 7; Page 585-586; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 1238; DB 3; Length 243;
Best Local Similarity 100.0%; Pred. No. 5e-90;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
```

RESULT 15
ABB73425
ID ABB73425 standard; protein; 243 AA.
XX
AC ABB73425;
XX
DT 05-APR-2002 (first entry)
XX
DE Fc-MMP inhibitor fusion nucleic acid SEQ ID NO:1067.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35775.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Example 7; Fig 25A-B; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 243 AA;

Query Match 100.0%; Score 1238; DB 5; Length 243;
Best Local Similarity 100.0%; Pred. No. 5e-90;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228


Search completed: April  4, 2006, 13:07:40
Job time : 115.105 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 37.3037 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-2
Perfect score: 1238
Sequence: 1 MDKTHTCPPCPAPELLGGPS.....MHEALHNHYTQKSLSLSPGK 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1233	99.6	255	4	S31866	Ig gamma-1 chain C
2	1233	99.6	330	1	GHHU	Ig gamma-1 chain C
3	1227	99.1	374	2	S69339	Ig heavy chain V r
4	1180	95.3	234	2	PT0207	Ig gamma chain C r
5	1146	92.6	377	2	A23511	Ig gamma-3 chain C
6	1144	92.4	377	2	A60764	Ig gamma-3 chain C
7	1142.5	92.3	326	1	G2HU	Ig gamma-2 chain C
8	1135	91.7	327	1	G4HU	Ig gamma-4 chain C
9	1121	90.5	289	1	G3HUWI	Ig gamma-3 heavy c
10	918.5	74.2	323	1	GHRB	Ig gamma chain C r
11	906.5	73.2	328	2	I47160	Ig gamma 2b chain
12	906.5	73.2	328	2	I47159	Ig gamma 2a chain
13	903.5	73.0	277	2	I47162	Ig gamma 4 chain c
14	889	71.8	329	1	G2GP	Ig gamma-2 chain C
15	885.5	71.5	328	2	I47158	Ig gamma 1 chain c
16	878.5	71.0	328	2	I47161	Ig gamma 3 chain c
17	855.5	69.1	470	2	S22080	Ig heavy chain pre
18	846	68.3	308	2	C30554	Ig heavy chain C r
19	846	68.3	472	2	S31459	Ig gamma-1 chain -
20	845.5	68.3	329	1	G3MSC	Ig gamma-3 chain C
21	838	67.7	333	2	PS0018	Ig gamma-2b chain
22	834.5	67.4	398	1	G3MSM	Ig gamma-3 chain C
23	827.5	66.8	444	2	PC4436	monoclonal antibod
24	818.5	66.1	326	2	PS0017	Ig gamma-1 chain C
25	817.5	66.0	324	1	G1MS	Ig gamma-1 chain C
26	812.5	65.6	393	1	G1MSM	Ig gamma-1 chain C
27	809.5	65.4	329	2	S00847	Ig gamma-2c chain
28	809	65.3	330	1	G2MSA	Ig gamma-2a chain
29	809	65.3	469	2	S37483	Ig gamma-2a chain
30	804	64.9	399	1	G2MSAM	Ig gamma-2a chain
31	802	64.8	335	1	G2MSAB	Ig gamma-2a chain
32	794	64.1	446	2	S40295	Ig gamma-2a chain
33	785.5	63.4	322	2	PS0019	Ig gamma-2a chain
34	779	62.9	474	1	G2MS11	Ig gamma-2b chain
35	774	62.5	405	1	G2MSBM	Ig gamma-2b chain
36	764	61.7	327	2	S06611	Ig gamma-2 chain C
37	757	61.1	475	2	S01321	Ig gamma-2b chain
38	707	57.1	180	2	I46732	Ig gamma heavy cha
39	577.5	46.6	249	2	S69340	Ig heavy chain VHI
40	574.5	46.4	218	2	A36040	Ig heavy chain V-I
41	571	46.1	152	2	S14236	Ig gamma-1 chain C
42	395.5	31.9	572	2	B46529	Ig Y heavy chain (
43	358	28.9	343	2	S25644	Ig mu chain C regi
44	358	28.9	453	2	S37768	Ig mu chain C regi
45	357.5	28.9	549	2	S04845	Ig heavy chain pre

ALIGNMENTS

RESULT 1
S31866
Ig gamma-1 chain C region - synthetic
C;Species: synthetic
A;Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C;Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C;Accession: S31866
R;Filpula, D.
submitted to the EMBL Data Library, February 1993
A;Description: Screeing method for protein-protein interactions of cloned gene producti
A;Reference number: S31866
A;Accession: S31866
A;Molecule type: mRNA
A;Residues: 1-255 <FIL>
A;Cross-references: UNIPARC:UPI000011F41F; EMBL:X70421; NID:g33068; PIDN:CAA49866.1; P
C;Keywords: immunoglobulin
F;1-22/Region: Escherichia coli outer membrane protein A precursor
F;23-255/Region: human Ig gamma-1 chain C region

Query Match 99.6%; Score 1233; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.7e-89;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 88

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 148

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 208

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 255
```

RESULT 2
GHHU
Ig gamma-1 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 09-Jul-2004
C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.
Nucleic Acids Res. 10, 4071-4079, 1982
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.
A;Reference number: A93433; MUID:82274238; PMID:6287432
A;Accession: A93433
A;Molecule type: DNA

A;Residues: 1-330 <ELL>
A;Cross-references: UNIPROT:P01857; UNIPARC:UPI0000034C0E; EMBL:Z17370
A;Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,
A;Note: Lys-330 is removed after translation
R;Harris, L.J.
submitted to the EMBL Data Library, October 1992
A;Reference number: S33904
A;Accession: S36861
A;Molecule type: DNA
A;Residues: 2-330 <HAR>
A;Cross-references: UNIPARC:UPI000013C6FE; EMBL:Z17370
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of a
A;Reference number: S33887; MUID:83001943; PMID:6811139
A;Accession: S33887
A;Molecule type: DNA
A;Residues: 88-113;235-330 <TAK>
A;Cross-references: UNIPARC:UPI000017378B; UNIPARC:UPI000017378C; EMBL:Z17370
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,
Biochemistry 9, 3161-3170, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen
A;Reference number: A90563; MUID:71064024; PMID:5489771
A;Contents: myeloma protein Eu
A;Accession: B90563
A;Molecule type: protein
A;Residues: 1-96,'R',98-135 <CUN>
A;Cross-references: UNIPARC:UPI000017378D
A;Note: this sequence has the G1m(3) marker, 97-Arg
R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid seque
A;Reference number: A90564; MUID:71064025; PMID:5530842
A;Contents: Eu
A;Accession: A90564
A;Molecule type: protein
A;Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',240,
A;Cross-references: UNIPARC:UPI000017378E
A;Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met
R;Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie),
igen Primaerstruktur.
A;Reference number: A91668; MUID:77070269; PMID:826475
A;Contents: myeloma protein Nie
A;Accession: B91668
A;Molecule type: protein
A;Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27
A;Cross-references: UNIPARC:UPI000017378F
A;Note: this sequence has the G1m(17) and G1m(1) markers
R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A;Title: Die Primaerstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1 KOL
A;Reference number: A91723; MUID:83289131; PMID:6884994
A;Contents: myeloma protein KOL; disulfide bonds
A;Accession: A91723
A;Molecule type: protein
A;Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH
A;Cross-references: UNIPARC:UPI0000173790
A;Note: this sequence has the G1m(3) and G1m(non-1) markers
R;Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
A;Reference number: A90565; MUID:71064027; PMID:4923144
A;Contents: annotation; disulfide bonds
R;Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A;Reference number: A91667; MUID:77070267; PMID:1002129
A;Contents: annotation; disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG1
A;Cross-references: GDB:120085; OMIM:147100
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 114/1; 224/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;137-206/Domain: immunoglobulin homology <IM2>
F;243-310/Domain: immunoglobulin homology <IM3>
F;27-83,144-204,250-308/Disulfide bonds: #status experimental
F;103/Disulfide bonds: interchain (to light chain) #status experimental
F;109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F;180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 99.6%; Score 1233; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 7.9e-89;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 223

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 283

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 330
```

RESULT 3
S69339
Ig heavy chain V region precursor - human
C;Species: Homo sapiens (man)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C;Accession: S69339; S72664
R;Khamlichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A;Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A;Reference number: S69339; MUID:95262687; PMID:7744049
A;Accession: S69339
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-374 <KHA>
A;Cross-references: UNIPARC:UPI0000176F24; EMBL:X81695
R;Khamlichi, A.A.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72664
A;Accession: S72664
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-140,'C',142-374 <KH2>
A;Cross-references: UNIPARC:UPI0000176F25; EMBL:X81695
C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 99.1%; Score 1227; DB 2; Length 374;
Best Local Similarity 99.1%; Pred. No. 2.7e-88;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 207

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 267
```

```
Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||:|:||||||||||||||||||||||||||||||||||||||||||
Db        268 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 327

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 374


RESULT 4
PT0207
Ig gamma chain C region - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C;Accession: PT0207
R;Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A;Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A;Reference number: PT0207; MUID:91287716; PMID:2062315
A;Accession: PT0207
A;Molecule type: mRNA
A;Residues: 1-234 <EHR>
A;Cross-references: UNIPARC:UPI0000176F05
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;48-117/Domain: immunoglobulin homology <IMM>

  Query Match             95.3%;  Score 1180;  DB 2;  Length 234;
  Best Local Similarity   98.6%;  Pred. No. 7e-85;
  Matches  217;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              | ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 DTTHTCPPCAAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 74

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 134

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        135 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYKTTPPVLDS 194

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 221
              |||||||||||||||||||||||||||||||||||||||||
Db        195 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 234


RESULT 5
A23511
Ig gamma-3 chain C region (allotype G3m(b)) - human
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
C;Accession: A23511
R;Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
Nucleic Acids Res. 14, 1779-1789, 1986
A;Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: co
A;Reference number: A23511; MUID:86148507; PMID:3081877
A;Accession: A23511
A;Molecule type: DNA
A;Residues: 1-377 <HUC>
A;Cross-references: UNIPARC:UPI000004718F; GB:X03604; GB:M12958; NID:g33070; PIDN:CAA272
C;Genetics:
A;Gene: GDB:IGHG3
A;Cross-references: GDB:119339; OMIM:147120
A;Map position: 14q32.33-14q32.33
A;Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;20-85/Domain: immunoglobulin homology <IMM>

  Query Match             92.6%;  Score 1146;  DB 2;  Length 377;
  Best Local Similarity   92.5%;  Pred. No. 5.7e-82;
  Matches  210;  Conservative    8;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |    || ||||||||||||||||||||||||||||||||||||||||||:| ||||
Db        151 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 210

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |
Db        211 GVEVHNAKTKPREEQYNSTFRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKTK 270

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||:|:||||||||||||||||||||:||||||| ||||:|||
Db        271 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDS 330

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||:|||||||||||| :|||||||||||
Db        331 DGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPGK 377


RESULT 6
A60764
Ig gamma-3 chain C region, form LAT - human
C;Species: Homo sapiens (man)
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 31-Dec-2004
C;Accession: A60764
R;Huck, S.; Lefranc, G.; Lefranc, M.P.
Immunogenetics 30, 250-257, 1989
A;Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conve
A;Reference number: A60764; MUID:90007613; PMID:2571587
A;Accession: A60764
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-377 <HUC>
A;Cross-references: UNIPROT:Q8N4Y9; UNIPARC:UPI0000176F0B
C;Superfamily: immunoglobulin homology
C;Keywords: immunoglobulin
F;20-85/Domain: immunoglobulin homology <IMM>

  Query Match             92.4%;  Score 1144;  DB 2;  Length 377;
  Best Local Similarity   92.5%;  Pred. No. 8.2e-82;
  Matches  210;  Conservative    8;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |    || ||||||||||||||||||||||||||||||||||||||||||:| ||||
Db        151 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 210

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |
Db        211 GVEVHNAKTKPREEQYNSTFRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKTK 270

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||:|:||||||||||||||||||||:||||||| ||||||||
Db        271 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPVLDS 330

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||:|||||||||:|||||||||||||||| :|||||||||||
Db        331 DGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNRFTQKSLSLSPGK 377


RESULT 7
G2HU
Ig gamma-2 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1981 #sequence_revision 13-Jun-1983 #text_change 09-Jul-2004
C;Accession: A93906; A92809; A90752; A93132; A02148
R;Ellison, J.; Hood, L.
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982
A;Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain c
A;Reference number: A93906; MUID:82197621; PMID:6804948
A;Accession: A93906
A;Molecule type: DNA
A;Residues: 1-326 <ELL>
A;Cross-references: UNIPROT:P01859; UNIPARC:UPI000003BFCC; GB:V00554; GB:J00230; NID:g
```

```
A;Note: Lys-326 is probably removed posttranslationally
R;Wang, A.C.; Tung, E.; Fudenberg, H.H.
J. Immunol. 125, 1048-1054, 1980
A;Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f
A;Reference number: A92809; MUID:81007873; PMID:6774012
A;Contents: myeloma protein Til
A;Accession: A92809
A;Molecule type: protein
A;Residues: 1-19,'Q',21-57,'Z',59,'A',61-193,'D',195-325 <WAN>
A;Cross-references: UNIPARC:UPI0000173791
A;Note: Trp-156 is at or near the complement-binding site
R;Connell, G.E.; Parr, D.M.; Hofmann, T.
Can. J. Biochem. 57, 758-767, 1979
A;Title: The amino acid sequences of the three heavy chain constant region domains of a
A;Reference number: A90752; MUID:80001357; PMID:113060
A;Contents: myeloma protein Zie
A;Accession: A90752
A;Molecule type: protein
A;Residues: 1-24,'E',26-57,'EV',60-85;132-171,'ZZZ',175,'B',177-193,'D',195-196,'Q',198-
A;Cross-references: UNIPARC:UPI0000173792; UNIPARC:UPI0000173793
A;Note: this sequence has since been revised
R;Hofmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A;Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g
A;Reference number: A93132; MUID:80114419; PMID:118920
A;Contents: Zie
A;Accession: A93132
A;Molecule type: protein
A;Residues: 238-275 <HOF>
A;Cross-references: UNIPARC:UPI0000173794
R;Hofmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A;Reference number: A94591
A;Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268
A;Note: the revised sequence differs from that shown in having 60-Ala and in the amidati
ned
R;Milstein, C.; Frangione, B.
Biochem. J. 121, 217-225, 1971
A;Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.
A;Reference number: A90253; MUID:72033500; PMID:4940472
A;Contents: annotation; myeloma protein Sa, disulfide bonds
R;Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A;Title: Structural studies of immunoglobulin G.
A;Reference number: A93157; MUID:69064124; PMID:5782707
A;Contents: annotation; Sa, disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG2
A;Cross-references: GDB:119338; OMIM:147110
A;Map position: 14q32.33-14q32.33
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;133-202/Domain: immunoglobulin homology <IM2>
F;239-306/Domain: immunoglobulin homology <IM3>
F;14/Disulfide bonds: interchain (to light chain) #status experimental
F;27-83,140-200,246-304/Disulfide bonds: #status experimental
F;102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F;176/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             92.3%;  Score 1142.5;  DB 1;  Length 326;
  Best Local Similarity   94.1%;  Pred. No. 8.9e-82;
  Matches  209;  Conservative    8;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          7 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
              |||||||  : |||||||||||||||||||||||||||||||||||||||:||||||||||
Db        106 CPPCPAPP-VAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVH 164

Qy         67 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 126
              ||||||||||:|||:||||||||:|||||||||||||||| |||||||||| ||||||
Db        165 NAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPRE 224

Qy        127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 186
              |||||||||:|:|||||||||||||||||||||||||||||||||||||:||||||||
Db        225 PQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFF 284

Qy        187 LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||
Db        285 LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 326


RESULT 8
G4HU
Ig gamma-4 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 09-Jul-2004
C;Accession: A90933; A90249; A02150
R;Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A;Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A;Reference number: A90933; MUID:83157104; PMID:6299662
A;Accession: A90933
A;Molecule type: DNA
A;Residues: 1-327 <ELL>
A;Cross-references: UNIPROT:P01861; UNIPARC:UPI0000047190
A;Note: the sequence was determined from the germline gene
R;Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A;Title: Human immunoglobulin sublclasses. Partial amino acid sequence of the constant
A;Reference number: A90249; MUID:70207560; PMID:4192699
A;Accession: A90249
A;Molecule type: protein
A;Residues: 1-30;81-326 <PIN>
A;Cross-references: UNIPARC:UPI0000173795; UNIPARC:UPI0000173796
C;Genetics:
A;Gene: GDB:IGHG4
A;Cross-references: GDB:119340; OMIM:147130
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 111/1; 221/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into l
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;99-110/Region: hinge
F;134-203/Domain: immunoglobulin homology <IM2>
F;240-307/Domain: immunoglobulin homology <IM3>
F;14/Disulfide bonds: interchain (to light chain) #status experimental
F;27-83,141-201,247-305/Disulfide bonds: #status predicted
F;106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F;177/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             91.7%;  Score 1135;  DB 1;  Length 327;
  Best Local Similarity   93.7%;  Pred. No. 3.4e-81;
  Matches  208;  Conservative    8;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          7 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
              || ||||| |||||||||||||||||||||||||||||||||| |||||:||||||||||
Db        106 CPSCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH 165

Qy         67 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 126
              ||||||||||:|||||||||||||||||||||||||||| ||: |||||||||||||
Db        166 NAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPRE 225

Qy        127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 186
              |||||||||::|:||||||||||||||||||||||||||||||||||||||||||||||
Db        226 PQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 285

Qy        187 LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||:|||||||||:|||||||||||||||||||||||| ||
Db        286 LYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLGK 327


RESULT 9
```

```
G3HUWI
Ig gamma-3 heavy chain disease proteins - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C;Accession: A90442; A92219; A90198; A93915; A02149
R;Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A;Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-
A;Reference number: A90442; MUID:81021548; PMID:6774747
A;Contents: heavy chain disease protein Wis
A;Accession: A90442
A;Molecule type: protein
A;Residues: 1-289 <FRA>
A;Cross-references: UNIPARC:UPI0000173797
A;Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
A;Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
A;Note: the sequence of residues 42-76 was taken from the reference that follows
R;Michaelsen, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A;Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
A;Reference number: A92219; MUID:77118561; PMID:402363
A;Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W
A;Accession: A92219
A;Molecule type: protein
A;Residues: 12-97 <MIC>
A;Cross-references: UNIPARC:UPI0000173798
A;Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
idue segment (12-28)
A;Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter
R;Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A;Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
A;Reference number: A90198; MUID:77021516; PMID:823945
A;Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues
A;Accession: A90198
A;Molecule type: protein
A;Residues: 59-125,'EB',128-226,228-289 <WOL>
A;Cross-references: UNIPARC:UPI0000173799
A;Note: this protein lacks most of the V region, all of the CH1 region, and part of the
R;Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A;Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
A;Reference number: A93915; MUID:82247835; PMID:6808505
A;Contents: heavy chain disease protein Omm
A;Accession: A93915
A;Molecule type: mRNA
A;Residues: 12-70;72-114;116-125,'E',127-133,'L',135-136,'E',138,'Y',140-154,'D',156-157
A;Cross-references: UNIPARC:UPI000017379A; UNIPARC:UPI000017379B; UNIPARC:UPI000017379C;
A;Note: a carboxyl-terminal Lys is removed posttranslationally
A;Note: this sequence may represent an allelic form or another gamma chain subclass
C;Comment: The heavy chain disease protein Wis is shown.
C;Genetics:
A;Gene: GDB:IGHG3
A;Cross-references: GDB:119339; OMIM:147120
A;Map position: 14q32.33-14q32.33
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
F;203-270/Domain: immunoglobulin homology <IMM>
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

  Query Match             90.5%;  Score 1121;  DB 1;  Length 289;
  Best Local Similarity   90.3%;  Pred. No. 3.6e-80;
  Matches  204;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |    || |||||||||||||||||||||||||||||||||||||||||||:| ||||
Db         64 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 123

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||:|||||||||:|:|||:||||||||||:||:|||||||||||||||||||||| |
Db        124 GVQVHNAKTKPREQQFNSTFRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTISKTK 183

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||:|:||||||||||||||||||||||||:||||||| ||||:|||
Db        184 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDS 243

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG 227
              ||||||||||||||||||||:|||||||||||| :||||||||||
Db        244 DGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPG 289


RESULT 10
GHRB
Ig gamma chain C region - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jul-2004
C;Accession: A91749; A90290; A93928; A90245; A94416; A02161
R;Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A;Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplo
A;Reference number: A91749; MUID:84030930; PMID:6313520
A;Accession: A91749
A;Molecule type: mRNA
A;Residues: 1-323 <BER>
A;Cross-references: UNIPROT:P01870; UNIPARC:UPI000012B37D
A;Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-T
R;Pratt, D.M.; Mole, L.E.
Biochem. J. 151, 337-349, 1975
A;Title: Sequence studies on the constant region of the Fd sections of rabbit immunogl
A;Reference number: A90290; MUID:76135469; PMID:1243651
A;Accession: A90290
A;Molecule type: protein
A;Residues: 1-47,'E',49-71,'PV',72-128 <PRA>
A;Cross-references: UNIPARC:UPI00001737AB
R;Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
A;Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy cha
A;Reference number: A93928; MUID:83299917; PMID:6193512
A;Accession: A93928
A;Molecule type: mRNA
A;Residues: 88-103,'M',105-143,'E',145-184,'A',186,'E',188-266 <MAR>
A;Cross-references: UNIPARC:UPI000016C5ED; GB:M16426; NID:g165111; PIDN:AAA31289.1; PI
A;Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic mar
R;Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
Biochem. J. 116, 249-259, 1970
A;Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobuli
A;Reference number: A90245; MUID:70110015; PMID:5461106
A;Accession: A90245
A;Molecule type: protein
A;Residues: 132-143,'E',145-161 <FRU>
A;Cross-references: UNIPARC:UPI00001737AC
R;Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiksel
A;Reference number: A94416
A;Accession: A94416
A;Molecule type: protein
A;Residues: 129-131;155-172,'D',174-184,'A',186,'E',188-200,'D',202-217,'E',219-232,'Q
A;Cross-references: UNIPARC:UPI00001737AD; UNIPARC:UPI00001737AE
A;Note: this has the e15 allotypic marker, 185-Ala
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (k
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-82/Domain: immunoglobulin homology <IM1>
F;130-199/Domain: immunoglobulin homology <IM2>
F;236-303/Domain: immunoglobulin homology <IM3>
F;173/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             74.2%;  Score 918.5;  DB 1;  Length 323;
  Best Local Similarity   71.7%;  Pred. No. 2.6e-64;
  Matches  167;  Conservative   29;  Mismatches   32;  Indels    5;  Gaps    2;

Qy          1 MDKT---HTC--PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 55
                :|||     ||  | || ||||||||||:|||||||||||||||||||||| :|||||:
Db         91 VDKTVAPSTCSKPTCPPPELLGGPSVFIFPPKPKDTLMISRTPEVTCVVVDVSQDDPEVQ 150
```

```
Qy          56 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 115
               | ||::  :|  |:   ||:|:||| |||| | : ||||| |||:|||| ||||||||||
Db         151 FTWYINNEQVRTARPPLREQQFNSTIRVVSTLPITHQDWLRGKEFKCKVHNKALPAPIEK 210

Qy         116 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 175
               |||||:||| ||:|||: | |:||:   |||||:: |||||||:|||| ||: |:|||||
Db         211 TISKARGQPLEPKVYTMGPPREELSSRSVSLTCMINGFYPSDISVEWEKNGKAEDNYKTT 270

Qy         176 PPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
               | |||||||:|||:||:|  | ||:|:|||:||||||||||||||||:| ||||
Db         271 PAVLDSDGSYFLYNKLSVPTSEWQRGDVFTCSVMHEALHNHYTQKSISRSPGK 323


RESULT 11
I47160
Ig gamma 2b chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47160
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47160
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115525; EMBL:U03780; NID:g433125; PIDN:AAA52218.1; PI
C;Genetics:
A;Gene: IgG2b
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain: immunoglobulin homology <IMM>

  Query Match             73.2%;  Score 906.5;  DB 2;  Length 328;
  Best Local Similarity   73.2%;  Pred. No. 2.3e-63;
  Matches  164;  Conservative   29;  Mismatches   28;  Indels    3;  Gaps    2;

Qy           7 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
               || ||| |   |||||:||||||||||||||:||||||||| |:|||:|:|||||||||
Db         106 CPICPACE-SPGPSVFIFPPKPKDTLMISRTPQVTCVVVDVSQENPEVQFSWYVDGVEVH 164

Qy          67 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 126
               |:|:|:|||:|||||||||| : |||||||||:||||:|| ||||| : ||||||| ||
Db         165 TAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITRIISKAKGQTRE 224

Qy         127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPPVLDSDGS 184
               ||||||||  :||::::||:|||| |||| || |||: |||  || ||:||||  | ||:
Db         225 PQVYTLPPHAEELSRSKVSITCLVIGFYPPDIDVEWQRNGQPEPEGNYRTTPPQQDVDGT 284

Qy         185 FFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
               :||||| :|||: || | :| |:|||||||||||:| :|||
Db         285 YFLYSKFSVDKASWQGGGIFQCAVMHEALHNHYTQKSISKTPGK 328


RESULT 12
I47159
Ig gamma 2a chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47159
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47159
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115524; EMBL:U03779; NID:g433123; PIDN:AAA52217.1; PI
C;Genetics:
A;Gene: IgG2a
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain: immunoglobulin homology <IMM>

  Query Match             73.2%;  Score 906.5;  DB 2;  Length 328;
  Best Local Similarity   73.2%;  Pred. No. 2.3e-63;
  Matches  164;  Conservative   29;  Mismatches   28;  Indels    3;  Gaps    2;

Qy           7 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
               || ||| |   |||||:||||||||||||||:||||||||| |:|||:|:|||||||||
Db         106 CPICPACE-SPGPSVFIFPPKPKDTLMISRTPQVTCVVVDVSQENPEVQFSWYVDGVEVH 164

Qy          67 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 126
               |:|:|:|||:|||||||||| : |||||||||:||||:|| ||||| : ||||||| ||
Db         165 TAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITRIISKAKGQTRE 224

Qy         127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPPVLDSDGS 184
               ||||||||  :||::::||:|||| |||| || |||: |||  || ||:||||  | ||:
Db         225 PQVYTLPPHAEELSRSKVSITCLVIGFYPPDIDVEWQRNGQPEPEGNYRTTPPQQDVDGT 284

Qy         185 FFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
               :||||| :|||: || | :| |:|||||||||||:| :|||
Db         285 YFLYSKFSVDKASWQGGGIFQCAVMHEALHNHYTQKSISKTPGK 328


RESULT 13
I47162
Ig gamma 4 chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47162
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47162
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-277 <KAC>
A;Cross-references: UNIPARC:UPI0000115527; EMBL:U03782; NID:g433129; PIDN:AAA52220.1; P
C;Genetics:
A;Gene: IgG4
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;82-151/Domain: immunoglobulin homology <IMM>

  Query Match             73.0%;  Score 903.5;  DB 2;  Length 277;
  Best Local Similarity   71.1%;  Pred. No. 3.2e-63;
  Matches  167;  Conservative   30;  Mismatches   31;  Indels    7;  Gaps    4;

Qy           1 MDK---THTCPPCP-APELLG-GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 55
               :||   | | ||||  |   | ||| |:||||||||||||||:||||||||| |:|||:
Db          43 VDKRVGTKTKPPCPICPACEGPGPSAFIFPPKPKDTLMISRTPKVTCVVVDVSQENPEVQ 102

Qy          56 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 115
               |:|||||||| |:|:|:|||:|||||||| : |||||||||:||||:|| ||||| :
Db         103 FSWYVDGVEVHTAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITR 162

Qy         116 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYK 173
               ||||||| |||||||||| :||::::|:||||| |||| || |||: |||  || ||:
Db         163 IISKAKGQTREPQVYTLPPPTEELSRSKVTLTCLVTGFYPPDIDVEWQRNGQPEPEGNYR 222

Qy         174 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
               |||| | ||::|||||| |||: ||:|: | |:|||||||||||||:  :|||
Db         223 TTPPQQDVDGTYFLYSKLAVDKASWQRGDTFQCAVMHEALHNHYTQKSIFKTPGK 277


RESULT 14
G2GP
Ig gamma-2 chain C region - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 09-Jul-2004
C;Accession: A94553; A90352; A90359; A90384; A90385; A02151
R;Trischmann, T.M.
submitted to the Atlas, April 1975
A;Reference number: A94553
```

A;Accession: A94553
A;Molecule type: protein
A;Residues: 1-3 <TRI>
A;Cross-references: UNIPROT:P01862; UNIPARC:UPI000017379E
R;Birshtein, B.K.; Hussain, Q.Z.; Cebra, J.J.
Biochemistry 10, 18-25, 1971
A;Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Am
A;Reference number: A90352; MUID:71058471; PMID:5538606
A;Accession: A90352
A;Molecule type: protein
A;Residues: 4-68 <BIR>
A;Cross-references: UNIPARC:UPI000017379F
R;Turner, K.J.; Cebra, J.J.
Biochemistry 10, 9-17, 1971
A;Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. Ami
A;Reference number: A90359; MUID:71058486; PMID:5538616
A;Accession: A90359
A;Molecule type: protein
A;Residues: 69-133;312-329 <TUR>
A;Cross-references: UNIPARC:UPI00001737A0; UNIPARC:UPI00001737A1
R;Tracey, D.E.; Cebra, J.J.
Biochemistry 13, 4796-4803, 1974
A;Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibodies.
A;Reference number: A90384; MUID:75036072; PMID:4429665
A;Accession: A90384
A;Molecule type: protein
A;Residues: 134-226 <TRA>
A;Cross-references: UNIPARC:UPI00001737A2
R;Trischmann, T.M.; Cebra, J.J.
Biochemistry 13, 4804-4811, 1974
A;Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibodies.
A;Reference number: A90385; MUID:75036073; PMID:4609467
A;Accession: A90385
A;Molecule type: protein
A;Residues: 227-311 <TR2>
A;Cross-references: UNIPARC:UPI00001737A3
R;Oliveira, B.; Lamm, M.E.
Biochemistry 10, 26-31, 1971
A;Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.
A;Reference number: A90354; MUID:71058474; PMID:4922544
A;Contents: annotation; disulfide bonds
A;Note: Cys-16 is involved in a heavy-light chain bond
A;Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds
C;Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;21-81/Domain: immunoglobulin homology <IM1>
F;135-204/Domain: immunoglobulin homology <IM2>
F;241-310/Domain: immunoglobulin homology <IM3>
F;28-79/Disulfide bonds: #status experimental
F;142-202/Disulfide bonds: #status experimental
F;178/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;248-308/Disulfide bonds: #status experimental

Query Match 71.8%; Score 889; DB 1; Length 329;
Best Local Similarity 72.3%; Pred. No. 5.3e-62;
Matches 162; Conservative 24; Mismatches 36; Indels 2; Gaps 1;

```
Qy          6 TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 65
              ||| || || |||||||:|||||||||||| || ||||||||| ::|||:| |:||   |
Db        106 TCPKCPPPENLGGPSVFIFPPKPKDTLMISLTPRVTCVVVDVSQDEPEVQFTWFVDNKPV 165

Qy         66 HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 125
               ||:|||| |||||:|:|| ||| : ||||| |||:|||| |||||||||||||| || ||
Db        166 GNAETKPRVEQYNTTFRVESVLPIQHQDWLRGKEFKCKVYNKALPAPIEKTISKTKGAPR 225

Qy        126 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP--ENNYKTTPPVLDSDG 183
               | |||||||||||:|::||:|||:  |:|:|| ||| ||  |  |  || |||: |:||
Db        226 MPDVYTLPPSRDELSKSKVSVTCLIINFFPADIHVEWASNRVPVSEKEYKNTPPIEDADG 285

Qy        184 SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG 227
```

```
                |:|||||||||||| | || |::|||||||||||| |||::| |||
Db        286 SYFLYSKLTVDKSAWDQGTVYTCSVMHEALHNHVTQKAISRSPG 329


RESULT 15
I47158
Ig gamma 1 chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47158
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47158
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115523; EMBL:U03778; NID:g433121; PIDN:AAA52216.1;
C;Genetics:
A;Gene: IgG1
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain:'immunoglobulin homology <IMM>

Query Match 71.5%; Score 885.5; DB 2; Length 328;
Best Local Similarity 72.4%; Pred. No. 1e-61;
Matches 163; Conservative 27; Mismatches 32; Indels 3; Gaps 2;

Qy          6 TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 65
              ||| ||  | : |||||:||||||||||||:||||||||||| |  ||:|:||||||||
Db        105 TCPICPGCE-VAGPSVFIFPPKPKDTLMISQTPEVTCVVVDVSKEHAEVQFSWYVDGVEV 163

Qy         66 HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 125
               | |:|:|:|||:|||||||||| : ||||| |||:||||:|  |||||  :|||||  || |
Db        164 HTAETRPKEEQFNSTYRVVSVLPIQHQDWLKGKEFKCKVNNVDLPAPITRTISKAIGQSR 223

Qy        126 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPPVLDSDG 183
              |||||||||  :||::::|:|||||  |||| || |||:||||  |||  |:||||  | ||
Db        224 EPQVYTLPPPAEELSRSKVTLTCLVIGFYPPDIHVEWKSNGQPEPENTYRTTPPQQDVDG 283

Qy        184 SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              :|||||| |||:||  |: | |:||||||||||||:| : ||
Db        284 TFFLYSKLAVDKARWDHGDKFECAVMHEALHNHYTQKSISKTQGK 328


Search completed: April 4, 2006, 13:17:23
Job time : 37.3037 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 174.283 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-2
Perfect score: 1238
Sequence: 1 MDKTHTCPPCPAPELLGGPS.....MHEALHNHYTQKSLSLSPGK 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1233	99.6	330	1	IGHG1_HUMAN	P01857 homo sapien
2	1233	99.6	465	2	Q6GMX6_HUMAN	Q6gmx6 homo sapien
3	1233	99.6	466	2	Q6IN78_HUMAN	Q6in78 homo sapien
4	1233	99.6	469	2	Q569F4_HUMAN	Q569f4 homo sapien
5	1233	99.6	469	2	Q7Z7P5_HUMAN	Q7z7p5 homo sapien
6	1233	99.6	470	2	Q7Z5W1_HUMAN	Q7z5w1 homo sapien
7	1233	99.6	470	2	Q6PJA4_HUMAN	Q6pja4 homo sapien
8	1233	99.6	472	2	Q6N089_HUMAN	Q6n089 homo sapien
9	1233	99.6	475	2	Q5EFE5_HUMAN	Q5efe5 homo sapien
10	1233	99.6	475	2	Q6GMW7_HUMAN	Q6gmw7 homo sapien
11	1233	99.6	476	2	Q6GMX1_HUMAN	Q6gmx1 homo sapien
12	1233	99.6	679	2	Q96PQ8_HUMAN	Q96pq8 homo sapien
13	1229	99.3	473	2	Q6P055_HUMAN	Q6p055 homo sapien
14	1229	99.3	475	2	Q6MZQ6_HUMAN	Q6mzq6 homo sapien
15	1229	99.3	480	2	Q6N094_HUMAN	Q6n094 homo sapien
16	1229	99.3	481	2	Q6N097_HUMAN	Q6n097 homo sapien
17	1229	99.3	482	2	Q7Z351_HUMAN	Q7z351 homo sapien
18	1227	99.1	348	2	Q6PYX1_HUMAN	Q6pyx1 homo sapien
19	1227	99.1	473	2	Q6MZV7_HUMAN	Q6mzv7 homo sapien
20	1227	99.1	478	2	Q6PI81_HUMAN	Q6pi81 homo sapien
21	1227	99.1	480	2	Q6PJF1_HUMAN	Q6pjf1 homo sapien
22	1226	99.0	466	2	Q6N096_HUMAN	Q6n096 homo sapien
23	1222	98.7	475	2	Q6N095_HUMAN	Q6n095 homo sapien
24	1222	98.7	544	2	Q6PJ95_HUMAN	Q6pj95 homo sapien
25	1216	98.2	487	2	Q65ZL2_9MURI	Q65zl2 mus sp. fv/
26	1172	94.7	475	2	Q5RE17_PONPY	Q5re17 pongo pygma
27	1146	92.6	354	2	Q86TT2_HUMAN	Q86tt2 homo sapien
28	1146	92.6	518	2	Q6N030_HUMAN	Q6n030 homo sapien
29	1146	92.6	519	2	Q5EBM2_HUMAN	Q5ebm2 homo sapien
30	1142.5	92.3	326	1	IGHG2_HUMAN	P01859 homo sapien
31	1142.5	92.3	417	2	Q6N093_HUMAN	Q6n093 homo sapien
32	1142	92.2	521	2	Q8N4Y9_HUMAN	Q8n4y9 homo sapien
33	1139.5	92.0	464	2	Q6MZU6_HUMAN	Q6mzu6 homo sapien
34	1137.5	91.9	465	2	Q6P6C4_HUMAN	Q6p6c4 homo sapien
35	1135	91.7	327	1	IGHG4_HUMAN	P01861 homo sapien
36	1135	91.7	473	2	Q8TC63_HUMAN	Q8tc63 homo sapien
37	1131	91.4	509	2	Q8NF17_HUMAN	Q8nf17 homo sapien
38	1128.5	91.2	470	2	Q68CN4_HUMAN	Q68cn4 homo sapien
39	1126	91.0	290	1	IGHG3_HUMAN	P01860 homo sapien
40	1126	91.0	476	2	Q6MZX7_HUMAN	Q6mzx7 homo sapien
41	918.5	74.2	323	1	GC_RABIT	P01870 oryctolagus
42	909	73.4	337	2	Q95M34_HORSE	Q95m34 equus cabal
43	889	71.8	329	1	IGHG2_CAVPO	P01862 cavia porce
44	845.5	68.3	329	1	GC3_MOUSE	P22436 mus musculu
45	845.5	68.3	470	2	Q7TMK1_MOUSE	Q7tmk1 mus musculu

ALIGNMENTS

RESULT 1
IGHG1_HUMAN
ID IGHG1_HUMAN STANDARD; PRT; 330 AA.
AC P01857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ig gamma-1 chain C region.
GN Name=IGHG1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=82274238; PubMed=6287432;
RA Ellison J.W., Berson B.J., Hood L.E.;
RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene.";
RL Nucleic Acids Res. 10:4071-4079(1982).
RN [2]
RP PROTEIN SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
RX MEDLINE=71064024; PubMed=5489771;
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA Waxdal M.J., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino
RT acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";
RL Biochemistry 9:3161-3170(1970).
RN [3]
RP PROTEIN SEQUENCE OF 136-329 (EU).
RX MEDLINE=71064025; PubMed=5530842;
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,
RA Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino
RT acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";
RL Biochemistry 9:3171-3181(1970).
RN [4]
RP PROTEIN SEQUENCE (MYELOMA PROTEIN NIE).
RX MEDLINE=77070269; PubMed=826475;
RA Ponstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic
RT peptides of the H-chain, alignment of the tryptic peptides and
RT discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [5]
RP PROTEIN SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [6]
RP DISULFIDE BONDS.

```
RX   MEDLINE=71064027; PubMed=4923144;
RA   Gall W.E., Edelman G.M.;
RT   "The covalent structure of a human gamma G-immunoglobulin. X.
RT   Intrachain disulfide bonds.";
RL   Biochemistry 9:3188-3196(1970).
RN   [7]
RP   DISULFIDE BONDS.
RX   MEDLINE=77070267; PubMed=1002129;
RA   Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT   "Rule of antibody structure. The primary structure of a monoclonal
RT   IgG1 immunoglobulin (myeloma protein Nie), I: purification and
RT   characterization of the protein, the L- and H-chains, the cyanogen
RT   bromide cleavage products, and the disulfide bridges.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX   MEDLINE=81208100; PubMed=7236608;
RA   Deisenhofer J.;
RT   "Crystallographic refinement and atomic models of a human Fc fragment
RT   and its complex with fragment B of protein A from Staphylococcus
RT   aureus at 2.9- and 2.8-A resolution.";
RL   Biochemistry 20:2361-2370(1981).
CC   -!- MISCELLANEOUS: Nie has the G1M(17) allotypic marker, 97-K, and the
CC       G1M(1) markers, 239-D and 241-L. KOL and EU sequences have the
CC       G1M(3) marker and the G1M (non-1) markers.
CC   -!- MISCELLANEOUS: Nie also differs in the amidation states of 35,
CC       116, 198, 269 and 272.
CC   -!- MISCELLANEOUS: EU also differs in the amidation states of residues
CC       155, 166, 177, 195, 198, 269, and 272 and in the order of residues
CC       268-272.
CC   -!- MISCELLANEOUS: KOL also differs in the amidation states of
CC       residues 198, 267 and 272.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; J00228; AAC82527.1; ALT_INIT; Genomic_DNA.
DR   PIR; A93433; GHHU.
DR   PDB; 1AJ7; X-ray; H=1-103.
DR   PDB; 1AQK; X-ray; H=1-103.
DR   PDB; 1D5B; X-ray; B/H=1-101.
DR   PDB; 1D5I; X-ray; H=1-101.
DR   PDB; 1D6V; X-ray; H=1-101.
DR   PDB; 1DN2; X-ray; A/B=120-326.
DR   PDB; 1E4K; X-ray; A/B=106-330.
DR   PDB; 1FC1; X-ray; A/B=106-329.
DR   PDB; 1FC2; X-ray; D=106-329.
DR   PDB; 1FCC; X-ray; A=121-326.
DR   PDB; 1HZH; X-ray; H/K=1-330.
DR   PDB; 1I7Z; X-ray; B/D=1-103.
DR   PDB; 1IIS; X-ray; A/B=107-330.
DR   PDB; 1IIX; X-ray; A/B=107-330.
DR   PDB; 1L6X; X-ray; A=120-326.
DR   PDB; 1OQX; X-ray; A/B=119-330.
DR   PDB; 1T83; X-ray; A/B=107-330.
DR   PDB; 2RCS; X-ray; H=1-103.
DR   HGNC; HGNC:5525; IGHG1.
DR   MIM; 147100; -.
DR   GO; GO:0005624; C:membrane fraction; NAS.
DR   GO; GO:0003823; F:antigen binding; TAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF07654; C1-set; 3.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS00290; IG_MHC; 2.
KW   3D-structure; Direct protein sequencing; Glycoprotein;
KW   Immunoglobulin C region; Immunoglobulin domain.
FT   REGION        1     98       CH1.
FT   REGION       99    110       Hinge.
FT   REGION      111    223       CH2.
FT   REGION      224    330       CH3.
FT   CARBOHYD    180    180       N-linked (GlcNAc. . .).
FT   DISULFID     27     83
FT   DISULFID    103    103       Interchain (with light chain).
FT   DISULFID    109    109       Interchain (with heavy chain).
FT   DISULFID    112    112       Interchain (with heavy chain).
FT   DISULFID    144    204
FT   DISULFID    250    308
FT   VARIANT      97     97       K -> R (in G1M(3) marker).
FT                                /FTId=VAR_003886.
FT   VARIANT     239    239       D -> E (in G1M(non-1) marker).
FT                                /FTId=VAR_003887.
FT   VARIANT     241    241       L -> M (in G1M(non-1) marker).
FT                                /FTId=VAR_003888.
FT   NON_TER       1      1
FT   STRAND       23     24
FT   STRAND       26     33
FT   STRAND       38     38
FT   STRAND       41     41
FT   TURN         42     45
FT   TURN         48     49
FT   STRAND       50     52
FT   STRAND       57     58
FT   TURN         59     61
FT   STRAND       62     71
FT   HELIX        73     75
FT   TURN         76     78
FT   STRAND       82     87
FT   TURN         88     91
FT   STRAND       92     97
FT   TURN        102    103
FT   STRAND      122    126
FT   HELIX       130    134
FT   TURN        136    137
FT   STRAND      141    149
FT   STRAND      157    162
FT   TURN        163    164
FT   STRAND      165    167
FT   STRAND      171    172
FT   STRAND      176    177
FT   TURN        179    180
FT   STRAND      183    190
FT   HELIX       193    197
FT   TURN        198    199
FT   STRAND      202    207
FT   TURN        209    210
FT   STRAND      215    219
FT   STRAND      227    227
FT   STRAND      230    234
FT   HELIX       238    242
FT   STRAND      245    256
FT   STRAND      261    266
FT   TURN        267    268
FT   STRAND      269    270
FT   STRAND      274    276
FT   STRAND      280    281
FT   TURN        283    284
FT   STRAND      287    296
FT   HELIX       297    301
FT   TURN        302    303
FT   STRAND      306    311
FT   TURN        313    314
FT   HELIX       316    318
FT   STRAND      319    324
SQ   SEQUENCE   330 AA;  36106 MW;  3770EE106C2FA33D CRC64;

  Query Match             99.6%;  Score 1233;  DB 1;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 223

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 283

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
                ||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 330


RESULT 2
Q6GMX6_HUMAN
ID   Q6GMX6_HUMAN PRELIMINARY;      PRT;   465 AA.
AC   Q6GMX6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RA   Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC073766; AAH73766.1; -; mRNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   465 AA;  51083 MW;  B3A9B7D0FDB1386E CRC64;
```

```
  Query Match           99.6%;  Score 1233;  DB 2;  Length 465;
  Best Local Similarity  100.0%;  Pred. No. 9e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 298

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 358

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 418

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
                ||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 465


RESULT 3
Q6IN78_HUMAN
ID   Q6IN78_HUMAN PRELIMINARY;      PRT;   466 AA.
AC   Q6IN78;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RG   NIH MGC Project;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072419; AAH72419.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
```

```
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   466 AA;  50854 MW;  53EB0BCEDE81076E CRC64;

  Query Match             99.6%;  Score 1233;  DB 2;  Length 466;
  Best Local Similarity   100.0%;  Pred. No. 9e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 299

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 359

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 419

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        420 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 466


RESULT 4
Q569F4_HUMAN
ID   Q569F4_HUMAN PRELIMINARY;      PRT;   469 AA.
AC   Q569F4;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lymph;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lymph;
RG   NIH MGC Project;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC092518; AAH92518.1; -; mRNA.
SQ   SEQUENCE   469 AA;  51254 MW;  AC13448E3047784F CRC64;

  Query Match             99.6%;  Score 1233;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 302

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 362

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 422

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        423 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 469


RESULT 5
Q7Z7P5_HUMAN
ID   Q7Z7P5_HUMAN PRELIMINARY;      PRT;   469 AA.
AC   Q7Z7P5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RG   NIH MGC Project;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC051328; AAH51328.1; -; mRNA.
DR   HSSP; P01857; 1HZH.
DR   SMR; Q7Z7P5; 20-469.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
```

```
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Immunoglobulin domain.
SQ   SEQUENCE   469 AA;  51395 MW;  C8D5BE12BAAF795C CRC64;

  Query Match             99.6%;  Score 1233;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 302

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 362

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 422

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        423 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 469


RESULT 6
Q7Z5W1_HUMAN
ID   Q7Z5W1_HUMAN PRELIMINARY;      PRT;   470 AA.
AC   Q7Z5W1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RA   Strausberg R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC053984; AAH53984.1; -; mRNA.
DR   HSSP; P01857; 1HZH.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   470 AA;  51204 MW;  778CF34521483E1A CRC64;

  Query Match             99.6%;  Score 1233;  DB 2;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 303

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 363

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 423

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        424 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 470


RESULT 7
Q6PJA4_HUMAN
ID   Q6PJA4_HUMAN PRELIMINARY;      PRT;   470 AA.
AC   Q6PJA4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RG   NIH MGC Project;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; BC018747; AAH18747.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   SMR; Q6PJA4; 20-470.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   470 AA;  51716 MW;  7B49556A11FD7D99 CRC64;

  Query Match             99.6%;  Score 1233;  DB 2;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 303

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 363

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 423

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        424 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 470


RESULT 8
Q6N089_HUMAN
ID   Q6N089_HUMAN PRELIMINARY;      PRT;   472 AA.
AC   Q6N089;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686P15220.
GN   Name=DKFZp686P15220;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Rectum tumor;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640627; CAE45781.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   472 AA;  51724 MW;  26CB340D0046D279 CRC64;

  Query Match             99.6%;  Score 1233;  DB 2;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 305

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 365

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 425

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        426 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 472


RESULT 9
Q5EFE5_HUMAN
ID   Q5EFE5_HUMAN PRELIMINARY;      PRT;   475 AA.
AC   Q5EFE5;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Anti-RhD monoclonal T125 gamma1 heavy chain precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Gaucher C., Klein P., Beliard R.;
RT   "Sequence determination of the recombinant human anti-RhD monoclonal
RT   antibody T125.";
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY894992; AAW82028.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   Pfam; PF07686; V-set; 1.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Signal.
FT   SIGNAL       1     19       Potential.
FT   CHAIN       20    475       anti-RhD monoclonal T125 gamma1 heavy
FT                               chain.
SQ   SEQUENCE   475 AA;  52362 MW;  1367D400DC7D2859 CRC64;

  Query Match             99.6%;  Score 1233;  DB 2;  Length 475;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
                ||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 475
```

RESULT 10
Q6GMW7_HUMAN
ID Q6GMW7_HUMAN PRELIMINARY; PRT; 475 AA.
AC Q6GMW7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC073782; AAH73782.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 475 AA; 51987 MW; 2A1FE55D736860F8 CRC64;

Query Match 99.6%; Score 1233; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 9.3e-92;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
                ||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 475
```

RESULT 11
Q6GMX1_HUMAN
ID Q6GMX1_HUMAN PRELIMINARY; PRT; 476 AA.
AC Q6GMX1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC073773; AAH73773.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 476 AA; 52286 MW; 622AABA5C62DDE9D CRC64;

Query Match 99.6%; Score 1233; DB 2; Length 476;

```
  Best Local Similarity   100.0%;  Pred. No. 9.3e-92;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 309

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 369

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 429

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 476


RESULT 12
Q96PQ8_HUMAN
ID   Q96PQ8_HUMAN PRELIMINARY;      PRT;   679 AA.
AC   Q96PQ8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VII active site mutant immunoconjugate.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21477448; PubMed=11593034; DOI=10.1073/pnas.201420298;
RA   Hu Z., Garen A.;
RT   "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT   cells for immunotherapy in mouse models of prostatic cancer.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Hu Z., Garen A.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF272774; AAK58686.2; -; mRNA.
DR   HSSP; P08709; 1KLI.
DR   SMR; Q96PQ8; 39-180, 191-444, 447-679.
DR   Ensembl; ENSG00000057593; Homo sapiens.
DR   GO; GO:0005576; C:extracellular region; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR001254; Peptidase_S1_S6.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF07654; C1-set; 2.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00594; Gla; 1.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00407; IGc1; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLA_1; UNKNOWN_1.
DR   PROSITE; PS50998; GLA_2; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
SQ   SEQUENCE   679 AA;  75552 MW;  0B0023AE70A067A1 CRC64;

  Query Match             99.6%;  Score 1233;  DB 2;  Length 679;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 512

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 572

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 632

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 679


RESULT 13
Q6P055_HUMAN
ID   Q6P055_HUMAN PRELIMINARY;      PRT;   473 AA.
AC   Q6P055;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
```

```
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RA   Strausberg R.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC065820; AAH65820.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   473 AA;  51344 MW;  9816D56A77129B57 CRC64;

  Query Match            99.3%;  Score 1229;  DB 2;  Length 473;
  Best Local Similarity  99.6%;  Pred. No. 1.9e-91;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 306

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 366

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 426

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||| ||||||||||||||||
Db        427 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 473


RESULT 14
Q6MZQ6_HUMAN
ID   Q6MZQ6_HUMAN PRELIMINARY;      PRT;   475 AA.
AC   Q6MZQ6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686G11190.
GN   Name=DKFZp686G11190;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Esophagus tumor;
RG   The German cDNA Consortium;
RA   Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA   Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640947; CAE45972.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   SMR; Q6MZQ6; 20-475.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   475 AA;  52043 MW;  B7EAE255A26F4B8E CRC64;

  Query Match            99.3%;  Score 1229;  DB 2;  Length 475;
  Best Local Similarity  99.6%;  Pred. No. 2e-91;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||| ||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 475


RESULT 15
Q6N094_HUMAN
ID   Q6N094_HUMAN PRELIMINARY;      PRT;   480 AA.
AC   Q6N094;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686O01196.
GN   Name=DKFZp686O01196;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Esophagus tumor;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640622; CAE45776.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   480 AA;  52612 MW;  225247F3D35AEC18 CRC64;

  Query Match            99.3%;  Score 1229;  DB 2;  Length 480;
  Best Local Similarity  99.6%;  Pred. No. 2e-91;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        254 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 313

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 373

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 433

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||| ||||||||||||||||
Db        434 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 480

Search completed: April  4, 2006, 13:15:14
Job time : 175.283 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 122.53 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-6
Perfect score: 1341
Sequence: 1 MDKTHTCPPCPAPELLGGPS.....KGGGGGIEGPTLRQWLAARA 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1341	100.0	247	3	AAB16958	Aab16958 Fc-TMP pr
2	1341	100.0	247	5	ABB73411	Abb73411 Fc-TPO mi
3	1341	100.0	268	3	AAB16959	Aab16959 Fc-TMP-TM
4	1341	100.0	268	5	ABB73412	Abb73412 Fc-TMP-TM
5	1341	100.0	269	3	AAY96531	Aay96531 Human IgG
6	1276	95.2	252	3	AAB17955	Aab17955 Fc-VEGF a
7	1276	95.2	252	5	ABB73423	Abb73423 Fc-VEGF a
8	1273	94.9	248	3	AAB17953	Aab17953 Fc-IL-1 a
9	1273	94.9	248	5	ABB73421	Abb73421 Fc-interl
10	1272.5	94.9	259	9	AEA18572	Aea18572 Amino aci
11	1270	94.7	282	5	AAU81169	Aau81169 Echistati
12	1269	94.6	243	7	ADN59746	Adn59746 Vector 20
13	1268	94.6	243	3	AAB17957	Aab17957 Fc-MMP in
14	1268	94.6	243	5	ABB73425	Abb73425 Fc-MMP in
15	1268	94.6	248	3	AAB17951	Aab17951 Fc-TNF-al
16	1268	94.6	248	5	ABB73419	Abb73419 Fc-TNF-al
17	1268	94.6	250	7	ADD31616	Add31616 Ang-2 pep
18	1268	94.6	250	8	ADT71978	Adt71978 Ang-2 neg
19	1268	94.6	253	3	AAB16964	Aab16964 Fc-EMP pr
20	1268	94.6	253	5	ABB73415	Abb73415 Fc-EPO mi
21	1268	94.6	277	3	AAB16967	Aab16967 Fc-EMP-EM
22	1268	94.6	277	5	ABB73418	Abb73418 Fc-EMP-EM
23	1266	94.4	436	7	ADC98616	Adc98616 Human ang
24	1264	94.3	271	5	AAE14335	Aae14335 Fc-calcit
25	1264	94.3	489	7	ADC98610	Adc98610 Human ang
26	1264	94.3	648	7	ADC98602	Adc98602 Human ang
27	1264	94.3	883	7	ADC98568	Adc98568 Human ang
28	1263	94.2	462	7	ADC98612	Adc98612 Human ang
29	1263	94.2	526	7	ADC98618	Adc98618 Human ang
30	1263	94.2	588	7	ADC98606	Adc98606 Human ang
31	1263	94.2	665	7	ADC98604	Adc98604 Human ang
32	1263	94.2	697	7	ADC98614	Adc98614 Human ang
33	1263	94.2	705	7	ADC98600	Adc98600 Human ang
34	1263	94.2	861	7	ADG76139	Adg76139 Human NOV
35	1263	94.2	861	8	ADF45362	Adf45362 Human sem
36	1263	94.2	861	8	ADO40296	Ado40296 Human sem
37	1263	94.2	878	7	ADG76141	Adg76141 Human NOV
38	1263	94.2	878	8	ADF45364	Adf45364 Human sem
39	1263	94.2	878	8	ADO40298	Ado40298 Human sem
40	1262.5	94.1	949	6	ABP70842	Abp70842 MVP-D pro
41	1258	93.8	465	9	ADX83744	Adx83744 Human IgG
42	1258	93.8	473	8	ADM97513	Adm97513 CD1d-IgG-
43	1258	93.8	485	9	ADX83727	Adx83727 Human IgG
44	1258	93.8	502	8	ADM97493	Adm97493 CD1d-IgG-
45	1258	93.8	588	7	ADC98608	Adc98608 Human ang

ALIGNMENTS

RESULT 1
AAB16958
ID AAB16958 standard; protein; 247 AA.
XX
AC AAB16958;
XX
DT 31-OCT-2000 (first entry)
XX
DE Fc-TMP protein sequence SEQ ID NO:6.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69444.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 21; Page 179-180; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 247 AA;

Query Match 100.0%; Score 1341; DB 3; Length 247;
Best Local Similarity 100.0%; Pred. No. 4.9e-94;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240

Qy        241 QWLAARA 247
              |||||||
Db        241 QWLAARA 247
```

RESULT 2
ABB73411
ID ABB73411 standard; protein; 247 AA.
XX
AC ABB73411;
XX
DT 05-APR-2002 (first entry)
XX
DE Fc-TPO mimetic peptide (Fc-TMP) amino acid SEQ ID NO:6.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35761.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 21; Fig 7; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 247 AA;

Query Match 100.0%; Score 1341; DB 5; Length 247;
Best Local Similarity 100.0%; Pred. No. 4.9e-94;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240

Qy        241 QWLAARA 247
              |||||||
Db        241 QWLAARA 247
```

RESULT 3
AAB16959
ID AAB16959 standard; protein; 268 AA.
XX
AC AAB16959;
XX
DT 31-OCT-2000 (first entry)
XX

```
DE   Fc-TMP-TMP protein sequence SEQ ID NO:8.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW   inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW   thrombosis; pharmaceutical.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-US025044.
XX
PR   23-OCT-1998;   98US-0105371P.
PR   22-OCT-1999;   99US-00428082.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham J,  Boone TC;
XX
DR   WPI; 2000-350702/30.
DR   N-PSDB; AAA69445.
XX
PT   Novel composition of matter comprising an Fc domain and pharmacologically
PT   active peptides, useful for treating cancer and autoimmune diseases.
XX
PS   Example 2; Page 182-183; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC   (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC   P3, and P4 = are each independently sequences of pharmacologically active
CC   peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC   c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC   of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC   thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC   cells from the present invention can be used for producing pharmaceutical
CC   compositions. The compositions are useful for treating cancer, asthma,
CC   thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC   a Fab domain) can provide a longer half-life or incorporate functions
CC   such as Fc receptor binding, protein A binding, complement fixation, and
CC   possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC   AAB18003 represent nucleotide and amino acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 268 AA;

  Query Match             100.0%;  Score 1341;  DB 3;  Length 268;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-94;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240

Qy        241 QWLAARA 247
              |||||||
Db        241 QWLAARA 247


RESULT 4
ABB73412
ID   ABB73412 standard; protein; 268 AA.
XX
AC   ABB73412;
XX
DT   05-APR-2002  (first entry)
XX
DE   Fc-TMP-TMP amino acid SEQ ID NO:8.
XX
KW   Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW   erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW   TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW   TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW   MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW   cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW   antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW   neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW   cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW   sleep disorder; neurological degenerative disease; anaemia;
KW   thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW   Fanconi's syndrome.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200183525-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014310.
XX
PR   03-MAY-2000; 2000US-00563286.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham JC,  Boone TC,  Gudas JM;
XX
DR   WPI; 2002-130313/17.
DR   N-PSDB; ABL35762.
XX
PT   Novel vehicle-peptide molecule or its multimers useful for treating
PT   inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT   diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS   Example 2; Fig 8; 176pp; English.
XX
CC   The present invention describes a vehicle-peptide molecule (I) or its
CC   multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC   cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC   antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC   neuroprotective activities. (I) can be used as a therapeutic or
CC   prophylactic agent as well as for screening purposes. (I) is useful for
CC   diagnosing diseases characterised by dysfunction of their associated
CC   protein of interest, for identifying normal or abnormal proteins of
CC   interest, as a part of diagnostic kit to detect the presence of their
CC   proteins of interest in a biological sample. Additionally, (I) is useful
CC   for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC   rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC   infertility, and neurological degenerative diseases. (I), comprising EPO-
CC   mimetic compounds are useful for treating disorders characterised by low
CC   red blood cell levels such as anaemia. The TPO-mimetic comprising
CC   compounds are useful for treating conditions that involve an existing
CC   megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC   deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC   tumour which result in thrombocytopaenia, systemic lupus erythematosus,
```

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 268 AA;

Query Match 100.0%; Score 1341; DB 5; Length 268;
Best Local Similarity 100.0%; Pred. No. 5.4e-94;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240

Qy        241 QWLAARA 247
              |||||||
Db        241 QWLAARA 247
```

RESULT 5
AAY96531
ID AAY96531 standard; protein; 269 AA.
XX
AC AAY96531;
XX
DT 04-SEP-2000 (first entry)
XX
DE Human IgG1 Fc TMP fusion protein.
XX
KW Immunoglobulin; IgG1; Fc; thrombopoietin; mimetic; TMP; TPO; platelet;
KW megakaryocyte; production; anti-human immunodeficiency virus; anti-HIV;
KW anti-anaemic; dermatological; immunosuppressive; anti-inflammatory.
XX
OS Homo sapiens.
XX
PN WO200024770-A2.
XX
PD 04-MAY-2000.
XX
PF 22-OCT-1999; 99WO-US024834.
XX
PR 23-OCT-1998; 98US-0105348P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Liu C, Feige U, Cheetham J;
XX
DR WPI; 2000-365108/31.
DR N-PSDB; AAA29229.
XX
PT Thrombopoietic peptides which activate mpl receptors and increase the
PT production of platelets or platelet precursors, useful for treatment of
PT diseases which involve thombocytopenia.
XX
PS Example 2A; Page 49-50; 91pp; English.
XX
CC A compound which binds to an mpl receptor comprising a thrombopoietin
CC mimetic peptide (TMP) dimer joined by a linker [TMP_1-(L_1)_nTMP_2], is
CC new. TMP_1 and TMP_2 are amino acid sequences varying from at least 10 to
CC 14 residues in length comprising X_2-X_1_0, X_2-X_1_1, X_2-X_1_2, X_2-
CC X_1_3, X_2-X_1_4, X_1-X_1_0, X_1-X_1_1, X_1-X_1_2, X_1-X_1_3, and X_1-
CC X_1_4. X_1 = I, A, V, L, S or R; X_2 = E, D, K or V; X_3 = G or A; X_4 =
CC P; X_5 = T or S; X_6 = L, I, V, A or F; X_7 = R or K; X_8 = Q, N, or E;
CC X_9 = W, Y or F; X_1_0 = L, I, V, A, F, M, or K; X_1_1 = A, I, V, L, F,
CC S, T, K, H, or E; X_1_2 = A, I, V, L, F, G, S, or Q; X_1_3 = R, K, T, V,
CC N, Q or G; X_1_4 = A, I, V, L, F, T, R, E, or G; L_1 = linker comprising
CC 1 to 20 amino acids; and n = 0 or 1. The compounds bind to and activate
CC the c-Mpl receptor which mediates the activity of endogenous
CC thrombopoietin. The TMPs are useful for increasing the production of
CC platelets or platelet precursors (e.g. megakaryocytes) in a mammal, which
CC is useful for treatment of diseases which involve thombocytopenia, e.g.
CC aplastic anaemia, immune thrombocytopenia (ITP), human immunodeficiency
CC virus associated ITP, and systemic lupus erythematosus
XX
SQ Sequence 269 AA;

Query Match 100.0%; Score 1341; DB 3; Length 269;
Best Local Similarity 100.0%; Pred. No. 5.5e-94;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240

Qy        241 QWLAARA 247
              |||||||
Db        241 QWLAARA 247
```

RESULT 6
AAB17955
ID AAB17955 standard; protein; 252 AA.
XX
AC AAB17955;
XX
DT 31-OCT-2000 (first entry)
XX
DE Fc-VEGF antagonist fusion protein sequence SEQ ID NO:1064.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX

DR WPI; 2000-350702/30.
DR N-PSDB; AAA69505.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 6; Page 579-580; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 252 AA;

Query Match 95.2%; Score 1276; DB 3; Length 252;
Best Local Similarity 99.6%; Pred. No. 4.5e-89;
Matches 234; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIE 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGVE 235
```

RESULT 7
ABB73423
ID ABB73423 standard; protein; 252 AA.
XX
AC ABB73423;
XX
DT 05-APR-2002 (first entry)
XX
DE Fc-VEGF antagonist fusion nucleic acid SEQ ID NO:1063.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35773.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Example 6; Fig 23A-B; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 252 AA;

Query Match 95.2%; Score 1276; DB 5; Length 252;
Best Local Similarity 99.6%; Pred. No. 4.5e-89;
Matches 234; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIE 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGVE 235
```

RESULT 8
AAB17953
ID AAB17953 standard; protein; 248 AA.

XX
AC AAB17953;
XX
DT 31-OCT-2000 (first entry)
XX
DE Fc-IL-1 antagonist fusion protein sequence SEQ ID NO:1060.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69503.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 5; Page 574-575; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 248 AA;

Query Match 94.9%; Score 1273; DB 3; Length 248;
Best Local Similarity 99.6%; Pred. No. 7.5e-89;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIE 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGFE 235
```

RESULT 9
ABB73421
ID ABB73421 standard; protein; 248 AA.
XX
AC ABB73421;
XX
DT 05-APR-2002 (first entry)
XX
DE Fc-interleukin 1 (IL-1) antagonist fusion nucleic acid SEQ ID NO:1059.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35771.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Example 5; Fig 21A-B; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,

CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 248 AA;

Query Match 94.9%; Score 1273; DB 5; Length 248;
Best Local Similarity 99.6%; Pred. No. 7.5e-89;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIE 235
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGFE 235
```

RESULT 10
AEA18572
ID AEA18572 standard; protein; 259 AA.
XX
AC AEA18572;
XX
DT 28-JUL-2005 (first entry)
XX
DE Amino acid sequence of a mL6-17 peptide fused to an Fc domain.
XX
KW immune reaction; immunogenic therapeutic agent; antibody titer; CTLA-4;
KW immunosuppressive; mL6-17; Fc domain.
XX
OS Synthetic.
XX
PN WO2005044188-A2.
XX
PD 19-MAY-2005.
XX
PF 26-OCT-2004; 2004WO-US035415.
XX
PR 27-OCT-2003; 2003US-0515199P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Khare SD, Feige U;
XX
DR WPI; 2005-346954/35.
XX
PT Decreasing immune reactions in a subject treated with a (potentially)
PT immunogenic therapeutic molecule comprises administering CTLA-4 within an
PT effective time interval relative to the administration of the
PT composition.
XX
PS Example 1; SEQ ID NO 6; 42pp; English.
XX
CC The specification describes a method of decreasing the incidence of an
CC immune reaction in a subject who is given a therapeutic composition
CC comprising a (potentially) immunogenic therapeutic molecule, tolerizing a
CC subject to such a molecule, or decreasing the antibody titer in a subject
CC administered such a molecule. The method comprises administering CTLA-4
CC to the subject within an effective time interval relative to the
CC administration of the therapeutic composition. The CTLA-4 may further
CC comprise an immunoglobulin heavy chain constant region. The method of the
CC invention is useful for modulating an immune response to an immunogenic
CC therapeutic agent. The present sequence represents a mL63-9 peptide fused
CC to an Fc domain. mL6-17 binds to nerve growth factor, and the fusion
CC protein is a therapeutic immunogenic molecule, which was used to
CC demonstrate the method of the invention.
XX
SQ Sequence 259 AA;

Query Match 94.9%; Score 1272.5; DB 9; Length 259;
Best Local Similarity 96.3%; Pred. No. 8.6e-89;
Matches 235; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIEGPTLR 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||| :   :
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGAQ---MI 237

Qy        241 QWLA 244
              ||:
Db        238 DWLS 241
```

RESULT 11
AAU81169
ID AAU81169 standard; protein; 282 AA.
XX
AC AAU81169;
XX
DT 29-AUG-2003 (revised)
DT 09-APR-2002 (first entry)
XX
DE Echistatin/IgG Fc fusion protein.
XX
KW IgG Fc; anticoagulant; thrombolytic; cytostatic; antiinflammatory;
KW immunosuppressive; osteopathic; antagonist; laminin; saw-scaled viper;
KW echistatin; integrin; selectin; vinculin; platelet aggregation;
KW angiogenesis; tumour; inflammation; autoimmune disease;
KW rheumatoid arthritis; osteoporosis.
XX
OS Echis carinatus.
OS Homo sapiens.
OS Chimeric.
XX
PN WO200181377-A2.
XX
PD 01-NOV-2001.
XX
PF 23-APR-2001; 2001WO-US013069.
XX
PR 21-APR-2000; 2000US-0198919P.
PR 03-MAY-2000; 2000US-0201394P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Kohno T, Lacey DL, Boone TC;
XX
DR WPI; 2002-062025/08.
DR N-PSDB; ABK24109.
XX
PT Composition comprising integrin or adhesion antagonistic peptide and
PT vehicle, useful for treating or preventing platelet aggregation, has a
PT longer half-life than free peptide.
XX
PS Example 1; Page 45-46; 68pp; English.

XX
CC The invention relates to a composition comprising an integrin/adhesion
CC antagonistic peptide (I) and a vehicle e.g. IgG Fc. The peptides are
CC based on laminin or saw-scaled viper echistatin and target integrin,
CC selectin or vinculin.Also included are compounds of formula (Ia) and
CC their multimers (X^1)_a-F^1-(X^2)_b where; F^1 = Fc domain; X^1 and X^2 =
CC -(L^1)_c-P^1, (L^1)_c-P^1-(L^2)_d-P^2, (L^1)_c-P^1-(L^2)_d-P^2-(L^3)_e-
CC P^3 or (L^1)_c-P^1-(L^2)_d-P^2-(L^3)_e-P^3-(L^4)_f-P^4; P^1-P^4 = same or
CC different (I); L^1-L^4 = same or different linkers; a-f = 0 or 1,
CC provided at least one of a and b = 1, a nucleic acid that encodes (Ia),
CC an expression vector containing the nucleic acid, host cells containing
CC the vector, producing a pharmaceutically active compound (B) by
CC covalently linking at least one Fc domain to at least one amino acid
CC sequence of a selected randomized (I) and any of six laminin-related
CC peptides (Ib). The compositions are used prophylactically and
CC therapeutically in the same way as (I), e.g. to inhibit platelet
CC aggregation or angiogenesis (tumours), or to treat inflammation and
CC autoimmune diseases (e.g. rheumatoid arthritis) and many different forms
CC of osteoporosis, also for diagnosis. Attaching the vehicle (especially Fc
CC domain) to (I) increases the half-life (free (I) are normally degraded
CC very quickly in vivo). The present sequence is a human IgG1 Fc-antagonist
CC peptide fusion compound of the invention. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
SQ Sequence 282 AA;

Query Match 94.7%; Score 1270; DB 5; Length 282;
Best Local Similarity 97.5%; Pred. No. 1.5e-88;
Matches 236; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGG--IEGPT 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||    ||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGECESGPC 240

Qy        239 LR 240
               |
Db        241 CR 242
```

RESULT 12
ADN59746
ID ADN59746 standard; protein; 243 AA.
XX
AC ADN59746;
XX
DT 01-JUL-2004 (first entry)
XX
DE Vector 20003182 encoded amino acid sequence, seq id 95.
XX
KW Haemostatic; antianaemic; immunosuppressive; platelet;
KW transmembrane signaling; mpl receptor; thrombopoietin mimetic peptide;
KW TMP; c-mpl receptor; platelet precursor; megakaryocyte;
KW thrombocytopaenia; aplastic anaemia; autoimmune thrombocytopaenia;
KW autoimmune haemolytic anaemia; Hughe's syndrome;
KW lupoid thrombocytopaenia.
XX
OS Unidentified.
XX
PN WO2003031589-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032552.
XX
PR 11-OCT-2001; 2001US-0328666P.
PR 10-OCT-2002; 2002US-00269806.
XX
PA (AMGE-) AMGEN INC.
XX
PI Min H, Sitney KC, Hartley C;
XX
DR WPI; 2003-403101/38.
DR N-PSDB; ADN59745.
XX
PT Novel thrombopoietin mimetic peptides which bind to mpl receptor, and
PT which stimulate the production of platelets and/or the production of
PT platelet precursors, useful for treating thrombocytopenia.
XX
PS Disclosure; SEQ ID NO 95; 126pp; English.
XX
CC The invention relates to a thrombopoietin mimetic peptide (TMP) (I) that
CC binds to the c-mpl (mpl) receptor, and which stimulates the production of
CC platelets and/or the production of platelet precursors, is new. Further
CC disclosed is a composition of matter (II) that binds to an mpl receptor,
CC and a pharmaceutical composition comprising (II) and a carrier. The
CC pharmaceutical composition of the invention is useful for treating
CC thrombocytopaenia in an animal, and for increasing megakaryocytes or
CC platelets in a patient. The TMP of the invention is useful for treating
CC conditions involving a megakaryocyte and/or platelet deficiency, e.g.
CC disease conditions involving thrombocytopaenia such as aplastic anaemia,
CC autoimmune thrombocytopaenia, drug induced immune thrombocytopaenia,
CC autoimmune haemolytic anaemia, Hughe's syndrome and lupoid
CC thrombocytopaenia. The TMP of the invention is also useful for
CC maintaining the viability or storage life of platelets and/or
CC megakaryocytes and its derived cells. The compounds demonstrate an
CC improved ability to bind to and/or trigger transmembrane signal through,
CC i.e. activating, the mpl receptor the compounds have superior
CC thromobopoietic activity, i.e. the ability to stimulate, in vivo and in
CC vitro, the production of platelets and/or megakaryocytopoietic activity,
CC i.e. the ability to stimulate, in vivo and in vitro, the production of
CC platelet precursors. Further, certain of the compounds also exhibit
CC superior therapeutic properties, such as improved plasma half-life,
CC biological activity and in vivo circulation time. The current sequence
CC represents the amino acid sequence encoded by a vector for use in
CC constructing C-terminal Fc fusion compounds (i.e. peptide attached at its
CC N-terminus to the C-terminus of the Fc).
XX
SQ Sequence 243 AA;

Query Match 94.6%; Score 1269; DB 7; Length 243;
Best Local Similarity 99.1%; Pred. No. 1.5e-88;
Matches 233; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGIE 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| :
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGAQ 235
```

RESULT 13
AAB17957
ID AAB17957 standard; protein; 243 AA.
XX

AC AAB17957;
XX
DT 31-OCT-2000 (first entry)
XX
DE Fc-MMP inhibitor fusion protein sequence SEQ ID NO:1068.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69507.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 7; Page 585-586; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 243 AA;

Query Match 94.6%; Score 1268; DB 3; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.8e-88;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGG 233
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGG 233
```

RESULT 14
ABB73425
ID ABB73425 standard; protein; 243 AA.
XX
AC ABB73425;
XX
DT 05-APR-2002 (first entry)
XX
DE Fc-MMP inhibitor fusion nucleic acid SEQ ID NO:1067.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35775.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Example 7; Fig 25A-B; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777

CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 243 AA;

Query Match 94.6%; Score 1268; DB 5; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.8e-88;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGG 233
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGG 233
```

RESULT 15
AAB17951
ID AAB17951 standard; protein; 248 AA.
XX
AC AAB17951;
XX
DT 31-OCT-2000 (first entry)
XX
DE Fc-TNF-alpha inhibitor fusion protein sequence SEQ ID NO:1056.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69501.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 4; Page 568-569; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 248 AA;

Query Match 94.6%; Score 1268; DB 3; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.8e-88;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

Qy         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120

Qy        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 180

Qy        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGG 233
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGG 233
```

Search completed: April 4, 2006, 13:07:44
Job time : 123.53 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 40.4123 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-6
Perfect score: 1341
Sequence: 1 MDKTHTCPPCPAPELLGGPS.....KGGGGGIEGPTLRQWLAARA 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1233	91.9	255	4	S31866	Ig gamma-1 chain C
2	1233	91.9	330	1	GHHU	Ig gamma-1 chain C
3	1227	91.5	374	2	S69339	Ig heavy chain V r
4	1180	88.0	234	2	PT0207	Ig gamma chain C r
5	1146	85.5	377	2	A23511	Ig gamma-3 chain C
6	1144	85.3	377	2	A60764	Ig gamma-3 chain C
7	1142.5	85.2	326	1	G2HU	Ig gamma-2 chain C
8	1135	84.6	327	1	G4HU	Ig gamma-4 chain C
9	1121	83.6	289	1	G3HUWI	Ig gamma-3 heavy c
10	918.5	68.5	323	1	GHRB	Ig gamma chain C r
11	906.5	67.6	328	2	I47160	Ig gamma 2b chain
12	906.5	67.6	328	2	I47159	Ig gamma 2a chain
13	903.5	67.4	277	2	I47162	Ig gamma 4 chain c
14	889	66.3	329	1	G2GP	Ig gamma-2 chain C
15	885.5	66.0	328	2	I47158	Ig gamma 1 chain c
16	878.5	65.5	328	2	I47161	Ig gamma 3 chain c
17	855.5	63.8	470	2	S22080	Ig heavy chain pre
18	846	63.1	308	2	C30554	Ig heavy chain C r
19	846	63.1	472	2	S31459	Ig gamma-1 chain -
20	845.5	63.0	329	1	G3MSC	Ig gamma-3 chain C
21	838	62.5	333	2	PS0018	Ig gamma-2b chain
22	834.5	62.2	398	1	G3MSM	Ig gamma-3 chain C
23	827.5	61.7	444	2	PC4436	monoclonal antibod
24	818.5	61.0	326	2	PS0017	Ig gamma-1 chain C
25	817.5	61.0	324	1	G1MS	Ig gamma-1 chain C
26	812.5	60.6	393	1	G1MSM	Ig gamma-1 chain C
27	809.5	60.4	329	2	S00847	Ig gamma-2c chain
28	809	60.3	330	1	G2MSA	Ig gamma-2a chain
29	809	60.3	469	2	S37483	Ig gamma-2a chain
30	804	60.0	399	1	G2MSAM	Ig gamma-2a chain
31	802	59.8	335	1	G2MSAB	Ig gamma-2a chain
32	794	59.2	446	2	S40295	Ig gamma-2a chain
33	785.5	58.6	322	2	PS0019	Ig gamma-2a chain
34	779	58.1	474	1	G2MS11	Ig gamma-2b chain
35	774	57.7	405	1	G2MSBM	Ig gamma-2b chain
36	764	57.0	327	2	S06611	Ig gamma-2 chain C
37	757	56.5	475	2	S01321	Ig gamma-2b chain
38	707	52.7	180	2	I46732	Ig gamma heavy cha
39	577.5	43.1	249	2	S69340	Ig heavy chain VHI
40	574.5	42.8	218	2	A36040	Ig heavy chain V-I
41	571	42.6	152	2	S14236	Ig gamma-1 chain C
42	395.5	29.5	572	2	B46529	Ig Y heavy chain (
43	360	26.8	343	2	S25644	Ig mu chain C regi
44	360	26.8	453	2	S37768	Ig mu chain C regi
45	359	26.8	455	1	MHMS	Ig mu chain C regi

ALIGNMENTS

RESULT 1
S31866
Ig gamma-1 chain C region - synthetic
C;Species: synthetic
A;Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C;Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C;Accession: S31866
R;Filpula, D.
submitted to the EMBL Data Library, February 1993
A;Description: Screeing method for protein-protein interactions of cloned gene product
A;Reference number: S31866
A;Accession: S31866
A;Molecule type: mRNA
A;Residues: 1-255 <FIL>
A;Cross-references: UNIPARC:UPI000011F41F; EMBL:X70421; NID:g33068; PIDN:CAA49866.1; P
C;Keywords: immunoglobulin
F;1-22/Region: Escherichia coli outer membrane protein A precursor
F;23-255/Region: human Ig gamma-1 chain C region

Query Match 91.9%; Score 1233; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 8.6e-88;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 88

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 148

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 208

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        209 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 255
```

RESULT 2
GHHU
Ig gamma-1 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 09-Jul-2004
C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.
Nucleic Acids Res. 10, 4071-4079, 1982
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.
A;Reference number: A93433; MUID:82274238; PMID:6287432
A;Accession: A93433
A;Molecule type: DNA

A;Residues: 1-330 <ELL>
A;Cross-references: UNIPROT:P01857; UNIPARC:UPI0000034C0E; EMBL:Z17370
A;Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,
A;Note: Lys-330 is removed after translation
R;Harris, L.J.
submitted to the EMBL Data Library, October 1992
A;Reference number: S33904
A;Accession: S36861
A;Molecule type: DNA
A;Residues: 2-330 <HAR>
A;Cross-references: UNIPARC:UPI000013C6FE; EMBL:Z17370
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of a
A;Reference number: S33887; MUID:83001943; PMID:6811139
A;Accession: S33887
A;Molecule type: DNA
A;Residues: 88-113;235-330 <TAK>
A;Cross-references: UNIPARC:UPI000017378B; UNIPARC:UPI000017378C; EMBL:Z17370
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,
Biochemistry 9, 3161-3170, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen
A;Reference number: A90563; MUID:71064024; PMID:5489771
A;Contents: myeloma protein Eu
A;Accession: B90563
A;Molecule type: protein
A;Residues: 1-96,'R',98-135 <CUN>
A;Cross-references: UNIPARC:UPI000017378D
A;Note: this sequence has the G1m(3) marker, 97-Arg
R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid seque
A;Reference number: A90564; MUID:71064025; PMID:5530842
A;Contents: Eu
A;Accession: A90564
A;Molecule type: protein
A;Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',240,
A;Cross-references: UNIPARC:UPI000017378E
A;Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met
R;Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie),
igen Primaerstruktur.
A;Reference number: A91668; MUID:77070269; PMID:826475
A;Contents: myeloma protein Nie
A;Accession: B91668
A;Molecule type: protein
A;Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27
A;Cross-references: UNIPARC:UPI000017378F
A;Note: this sequence has the G1m(17) and G1m(1) markers
R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A;Title: Die Primaerstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1 KOL
A;Reference number: A91723; MUID:83289131; PMID:6884994
A;Contents: myeloma protein KOL; disulfide bonds
A;Accession: A91723
A;Molecule type: protein
A;Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH
A;Cross-references: UNIPARC:UPI0000173790
A;Note: this sequence has the G1m(3) and G1m(non-1) markers
R;Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
A;Reference number: A90565; MUID:71064027; PMID:4923144
A;Contents: annotation; disulfide bonds
R;Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A;Reference number: A91667; MUID:77070267; PMID:1002129
A;Contents: annotation; disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG1
A;Cross-references: GDB:120085; OMIM:147100
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 114/1; 224/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into l
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;137-206/Domain: immunoglobulin homology <IM2>
F;243-310/Domain: immunoglobulin homology <IM3>
F;27-83,144-204,250-308/Disulfide bonds: #status experimental
F;103/Disulfide bonds: interchain (to light chain) #status experimental
F;109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F;180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 91.9%; Score 1233; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.2e-87;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 223

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 283

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        284 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 330
```

RESULT 3
S69339
Ig heavy chain V region precursor - human
C;Species: Homo sapiens (man)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C;Accession: S69339; S72664
R;Khamlichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A;Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A;Reference number: S69339; MUID:95262687; PMID:7744049
A;Accession: S69339
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-374 <KHA>
A;Cross-references: UNIPARC:UPI0000176F24; EMBL:X81695
R;Khamlichi, A.A.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72664
A;Accession: S72664
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-140,'C',142-374 <KH2>
A;Cross-references: UNIPARC:UPI0000176F25; EMBL:X81695
C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 91.5%; Score 1227; DB 2; Length 374;
Best Local Similarity 99.1%; Pred. No. 4e-87;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 207

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 267
```

```
Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||:|:||||||||||||||||||||||||||||||||||||||||||
Db        268 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 327

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 374
```

RESULT 4
PT0207
Ig gamma chain C region - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C;Accession: PT0207
R;Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A;Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A;Reference number: PT0207; MUID:91287716; PMID:2062315
A;Accession: PT0207
A;Molecule type: mRNA
A;Residues: 1-234 <EHR>
A;Cross-references: UNIPARC:UPI0000176F05
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;48-117/Domain: immunoglobulin homology <IMM>

```
  Query Match             88.0%;  Score 1180;  DB 2;  Length 234;
  Best Local Similarity   98.6%;  Pred. No. 9.3e-84;
  Matches  217;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              | ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 DTTHTCPPCAAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 74

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 134

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        135 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYKTTPPVLDS 194

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 221
              |||||||||||||||||||||||||||||||||||||||||
Db        195 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 234
```

RESULT 5
A23511
Ig gamma-3 chain C region (allotype G3m(b)) - human
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
C;Accession: A23511
R;Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
Nucleic Acids Res. 14, 1779-1789, 1986
A;Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: cc
A;Reference number: A23511; MUID:86148507; PMID:3081877
A;Accession: A23511
A;Molecule type: DNA
A;Residues: 1-377 <HUC>
A;Cross-references: UNIPARC:UPI000004718F; GB:X03604; GB:M12958; NID:g33070; PIDN:CAA272
C;Genetics:
A;Gene: GDB:IGHG3
A;Cross-references: GDB:119339; OMIM:147120
A;Map position: 14q32.33-14q32.33
A;Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;20-85/Domain: immunoglobulin homology <IMM>

```
  Query Match             85.5%;  Score 1146;  DB 2;  Length 377;
  Best Local Similarity   92.5%;  Pred. No. 7.1e-81;
  Matches  210;  Conservative    8;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |    || ||||||||||||||||||||||||||||||||||||||||||||:| ||||
Db        151 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 210

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||| |
Db        211 GVEVHNAKTKPREEQYNSTFRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKTK 270

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||:|:||||||||||||||||||||||:||||||| ||||:|||
Db        271 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDS 330

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||:|||||||||||| :|||||||||||
Db        331 DGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPGK 377
```

RESULT 6
A60764
Ig gamma-3 chain C region, form LAT - human
C;Species: Homo sapiens (man)
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 31-Dec-2004
C;Accession: A60764
R;Huck, S.; Lefranc, G.; Lefranc, M.P.
Immunogenetics 30, 250-257, 1989
A;Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conve
A;Reference number: A60764; MUID:90007613; PMID:2571587
A;Accession: A60764
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-377 <HUC>
A;Cross-references: UNIPROT:Q8N4Y9; UNIPARC:UPI0000176F0B
C;Superfamily: immunoglobulin homology
C;Keywords: immunoglobulin
F;20-85/Domain: immunoglobulin homology <IMM>

```
  Query Match             85.3%;  Score 1144;  DB 2;  Length 377;
  Best Local Similarity   92.5%;  Pred. No. 1e-80;
  Matches  210;  Conservative    8;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |    || ||||||||||||||||||||||||||||||||||||||||||||:| ||||
Db        151 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 210

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||| |
Db        211 GVEVHNAKTKPREEQYNSTFRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKTK 270

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||:|:||||||||||||||||||||||:||||||| ||||||||
Db        271 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPVLDS 330

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||:|||||||||:|||||||||||||||| :||||||||||
Db        331 DGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNRFTQKSLSLSPGK 377
```

RESULT 7
G2HU
Ig gamma-2 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1981 #sequence_revision 13-Jun-1983 #text_change 09-Jul-2004
C;Accession: A93906; A92809; A90752; A93132; A02148
R;Ellison, J.; Hood, L.
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982
A;Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain cc
A;Reference number: A93906; MUID:82197621; PMID:6804948
A;Accession: A93906
A;Molecule type: DNA
A;Residues: 1-326 <ELL>
A;Cross-references: UNIPROT:P01859; UNIPARC:UPI000003BFCC; GB:V00554; GB:J00230; NID:g3

A;Note: Lys-326 is probably removed posttranslationally
R;Wang, A.C.; Tung, E.; Fudenberg, H.H.
J. Immunol. 125, 1048-1054, 1980
A;Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f
A;Reference number: A92809; MUID:81007873; PMID:6774012
A;Contents: myeloma protein Til
A;Accession: A92809
A;Molecule type: protein
A;Residues: 1-19,'Q',21-57,'Z',59,'A',61-193,'D',195-325 <WAN>
A;Cross-references: UNIPARC:UPI0000173791
A;Note: Trp-156 is at or near the complement-binding site
R;Connell, G.E.; Parr, D.M.; Hofmann, T.
Can. J. Biochem. 57, 758-767, 1979
A;Title: The amino acid sequences of the three heavy chain constant region domains of a
A;Reference number: A90752; MUID:80001357; PMID:113060
A;Contents: myeloma protein Zie
A;Accession: A90752
A;Molecule type: protein
A;Residues: 1-24,'E',26-57,'EV',60-85;132-171,'ZZZ',175,'B',177-193,'D',195-196,'Q',198-
A;Cross-references: UNIPARC:UPI0000173792; UNIPARC:UPI0000173793
A;Note: this sequence has since been revised
R;Hofmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A;Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g
A;Reference number: A93132; MUID:80114419; PMID:118920
A;Contents: Zie
A;Accession: A93132
A;Molecule type: protein
A;Residues: 238-275 <HOF>
A;Cross-references: UNIPARC:UPI0000173794
R;Hofmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A;Reference number: A94591
A;Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268
A;Note: the revised sequence differs from that shown in having 60-Ala and in the amidati
ned
R;Milstein, C.; Frangione, B.
Biochem. J. 121, 217-225, 1971
A;Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.
A;Reference number: A90253; MUID:72033500; PMID:4940472
A;Contents: annotation; myeloma protein Sa, disulfide bonds
R;Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A;Title: Structural studies of immunoglobulin G.
A;Reference number: A93157; MUID:69064124; PMID:5782707
A;Contents: annotation; Sa, disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG2
A;Cross-references: GDB:119338; OMIM:147110
A;Map position: 14q32.33-14q32.33
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;133-202/Domain: immunoglobulin homology <IM2>
F;239-306/Domain: immunoglobulin homology <IM3>
F;14/Disulfide bonds: interchain (to light chain) #status experimental
F;27-83,140-200,246-304/Disulfide bonds: #status experimental
F;102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F;176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.2%; Score 1142.5; DB 1; Length 326;
Best Local Similarity 94.1%; Pred. No. 1.1e-80;
Matches 209; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

```
Qy          7 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
              ||||||| : |||||||||||||||||||||||||||||||||||||:||||||||||
Db        106 CPPCPAPP-VAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVH 164

Qy         67 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 126
              ||||||||||:|||:||||||||:|||||||||||||||| |||||||||| ||||||
Db        165 NAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPRE 224

Qy        127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 186
              |||||||||:|:||||||||||||||||||||||||||||||||||||||:||||||||
Db        225 PQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFF 284

Qy        187 LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||
Db        285 LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 326
```

RESULT 8
G4HU
Ig gamma-4 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 09-Jul-2004
C;Accession: A90933; A90249; A02150
R;Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A;Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A;Reference number: A90933; MUID:83157104; PMID:6299662
A;Accession: A90933
A;Molecule type: DNA
A;Residues: 1-327 <ELL>
A;Cross-references: UNIPROT:P01861; UNIPARC:UPI0000047190
A;Note: the sequence was determined from the germline gene
R;Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A;Title: Human immunoglobulin sublclasses. Partial amino acid sequence of the constant
A;Reference number: A90249; MUID:70207560; PMID:4192699
A;Accession: A90249
A;Molecule type: protein
A;Residues: 1-30;81-326 <PIN>
A;Cross-references: UNIPARC:UPI0000173795; UNIPARC:UPI0000173796
C;Genetics:
A;Gene: GDB:IGHG4
A;Cross-references: GDB:119340; OMIM:147130
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 111/1; 221/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into l
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;99-110/Region: hinge
F;134-203/Domain: immunoglobulin homology <IM2>
F;240-307/Domain: immunoglobulin homology <IM3>
F;14/Disulfide bonds: interchain (to light chain) #status experimental
F;27-83,141-201,247-305/Disulfide bonds: #status predicted
F;106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F;177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.6%; Score 1135; DB 1; Length 327;
Best Local Similarity 93.7%; Pred. No. 4.2e-80;
Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

```
Qy          7 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
              || ||||| |||||||||||||||||||||||||||||||| ||||||:||||||||||
Db        106 CPSCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVH 165

Qy         67 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 126
              ||||||||||:|||||||||||||||||||||||||||||| ||: |||||||||||||
Db        166 NAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPRE 225

Qy        127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 186
              |||||||||::|:|||||||||||||||||||||||||||||||||||||||||||||
Db        226 PQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 285

Qy        187 LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||:|||||||||:|||||||||||||||||||||||||| ||
Db        286 LYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLSLSLGK 327
```

RESULT 9

```
G3HUWI
Ig gamma-3 heavy chain disease proteins - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C;Accession: A90442; A92219; A90198; A93915; A02149
R;Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A;Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-
A;Reference number: A90442; MUID:81021548; PMID:6774747
A;Contents: heavy chain disease protein Wis
A;Accession: A90442
A;Molecule type: protein
A;Residues: 1-289 <FRA>
A;Cross-references: UNIPARC:UPI0000173797
A;Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
A;Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
A;Note: the sequence of residues 42-76 was taken from the reference that follows
R;Michaelsen, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A;Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
A;Reference number: A92219; MUID:77118561; PMID:402363
A;Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W
A;Accession: A92219
A;Molecule type: protein
A;Residues: 12-97 <MIC>
A;Cross-references: UNIPARC:UPI0000173798
A;Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
idue segment (12-28)
A;Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter
R;Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A;Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
A;Reference number: A90198; MUID:77021516; PMID:823945
A;Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues
A;Accession: A90198
A;Molecule type: protein
A;Residues: 59-125,'EB',128-226,228-289 <WOL>
A;Cross-references: UNIPARC:UPI0000173799
A;Note: this protein lacks most of the V region, all of the CH1 region, and part of the
R;Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A;Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
A;Reference number: A93915; MUID:82247835; PMID:6808505
A;Contents: heavy chain disease protein Omm
A;Accession: A93915
A;Molecule type: mRNA
A;Residues: 12-70;72-114;116-125,'E',127-133,'L',135-136,'E',138,'Y',140-154,'D',156-157
A;Cross-references: UNIPARC:UPI000017379A; UNIPARC:UPI000017379B; UNIPARC:UPI000017379C;
A;Note: a carboxyl-terminal Lys is removed posttranslationally
A;Note: this sequence may represent an allelic form or another gamma chain subclass
C;Comment: The heavy chain disease protein Wis is shown.
C;Genetics:
A;Gene: GDB:IGHG3
A;Cross-references: GDB:119339; OMIM:147120
A;Map position: 14q32.33-14q32.33
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
F;203-270/Domain: immunoglobulin homology <IMM>
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

  Query Match             83.6%;  Score 1121;  DB 1;  Length 289;
  Best Local Similarity   90.3%;  Pred. No. 4.3e-79;
  Matches  204;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |    || |||||||||||||||||||||||||||||||||||||||||||||:| ||||
Db         64 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 123

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||:||||||||||:|:|||:||||||||||:||:|||||||||||||||||||||| |
Db        124 GVQVHNAKTKPREQQFNSTFRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTISKTK 183

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||:|:|||||||||||||||||||||||||:||||||| ||||:|||
Db        184 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDS 243

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG 227
              |||||||||||||||||||:|||||||||||| :||||||||||
Db        244 DGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPG 289


RESULT 10
GHRB
Ig gamma chain C region - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jul-2004
C;Accession: A91749; A90290; A93928; A90245; A94416; A02161
R;Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A;Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplo
A;Reference number: A91749; MUID:84030930; PMID:6313520
A;Accession: A91749
A;Molecule type: mRNA
A;Residues: 1-323 <BER>
A;Cross-references: UNIPROT:P01870; UNIPARC:UPI000012B37D
A;Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Th
R;Pratt, D.M.; Mole, L.E.
Biochem. J. 151, 337-349, 1975
A;Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglo
A;Reference number: A90290; MUID:76135469; PMID:1243651
A;Accession: A90290
A;Molecule type: protein
A;Residues: 1-47,'E',49-71,'PV',72-128 <PRA>
A;Cross-references: UNIPARC:UPI00001737AB
R;Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
A;Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy chai
A;Reference number: A93928; MUID:83299917; PMID:6193512
A;Accession: A93928
A;Molecule type: mRNA
A;Residues: 88-103,'M',105-143,'E',145-184,'A',186,'E',188-266 <MAR>
A;Cross-references: UNIPARC:UPI000016C5ED; GB:M16426; NID:g165111; PIDN:AAA31289.1; PID
A;Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic mark
R;Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
Biochem. J. 116, 249-259, 1970
A;Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin
A;Reference number: A90245; MUID:70110015; PMID:5461106
A;Accession: A90245
A;Molecule type: protein
A;Residues: 132-143,'E',145-161 <FRU>
A;Cross-references: UNIPARC:UPI00001737AC
R;Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiksell
A;Reference number: A94416
A;Accession: A94416
A;Molecule type: protein
A;Residues: 129-131;155-172,'D',174-184,'A',186,'E',188-200,'D',202-217,'E',219-232,'Q'
A;Cross-references: UNIPARC:UPI00001737AD; UNIPARC:UPI00001737AE
A;Note: this has the e15 allotypic marker, 185-Ala
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into l
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-82/Domain: immunoglobulin homology <IM1>
F;130-199/Domain: immunoglobulin homology <IM2>
F;236-303/Domain: immunoglobulin homology <IM3>
F;173/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             68.5%;  Score 918.5;  DB 1;  Length 323;
  Best Local Similarity   71.7%;  Pred. No. 2e-63;
  Matches  167;  Conservative   29;  Mismatches   32;  Indels    5;  Gaps    2;

Qy          1 MDKT---HTC--PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 55
                :|||    ||  | || ||||||||||:||||||||||||||||||||| :|||||:
Db         91 VDKTVAPSTCSKPTCPPPELLGGPSVFIFPPKPKDTLMISRTPEVTCVVVDVSQDDPEVQ 150
```

```
Qy          56 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 115
               | ||::  :|  |:   ||:|:||| |||| | : ||||| |||:|||| ||||||||||
Db         151 FTWYINNEQVRTARPPLREQQFNSTIRVVSTLPITHQDWLRGKEFKCKVHNKALPAPIEK 210

Qy         116 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 175
               |||||:||| ||:|||: | |:||:   |||||:: ||||||:|||| ||: |:|||||
Db         211 TISKARGQPLEPKVYTMGPPREELSSRSVSLTCMINGFYPSDISVEWEKNGKAEDNYKTT 270

Qy         176 PPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
               | |||||||:|||:||:|  | ||:|:||:||||||||||||||||:| ||||
Db         271 PAVLDSDGSYFLYNKLSVPTSEWQRGDVFTCSVMHEALHNHYTQKSISRSPGK 323


RESULT 11
I47160
Ig gamma 2b chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47160
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47160
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115525; EMBL:U03780; NID:g433125; PIDN:AAA52218.1; PI
C;Genetics:
A;Gene: IgG2b
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain: immunoglobulin homology <IMM>

  Query Match             67.6%;  Score 906.5;  DB 2;  Length 328;
  Best Local Similarity   73.2%;  Pred. No. 1.7e-62;
  Matches  164;  Conservative   29;  Mismatches   28;  Indels    3;  Gaps    2;

Qy           7 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
               || ||| |   |||||:||||||||||||||:||||||||| |:|||:|:|||||||||
Db         106 CPICPACE-SPGPSVFIFPPKPKDTLMISRTPQVTCVVVDVSQENPEVQFSWYVDGVEVH 164

Qy          67 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 126
               |:|:|:|||:|||||||||| : |||||||||:||||:|| ||||| : ||||||| ||
Db         165 TAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITRIISKAKGQTRE 224

Qy         127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPPVLDSDGS 184
               ||||||||  :||::::||:|||| |||| || |||: |||  || ||:||||  | ||:
Db         225 PQVYTLPPHAEELSRSKVSITCLVIGFYPPDIDVEWQRNGQPEPEGNYRTTPPQQDVDGT 284

Qy         185 FFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
               :||||| :|||: || | :| |:||||||||||||:| :|||
Db         285 YFLYSKFSVDKASWQGGGIFQCAVMHEALHNHYTQKSISKTPGK 328


RESULT 12
I47159
Ig gamma 2a chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47159
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47159
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115524; EMBL:U03779; NID:g433123; PIDN:AAA52217.1; PI
C;Genetics:
A;Gene: IgG2a
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain: immunoglobulin homology <IMM>

  Query Match             67.6%;  Score 906.5;  DB 2;  Length 328;
  Best Local Similarity   73.2%;  Pred. No. 1.7e-62;
  Matches  164;  Conservative   29;  Mismatches   28;  Indels    3;  Gaps    2;

Qy           7 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66
               || ||| |   |||||:||||||||||||||:||||||||| |:|||:|:|||||||||
Db         106 CPICPACE-SPGPSVFIFPPKPKDTLMISRTPQVTCVVVDVSQENPEVQFSWYVDGVEVH 164

Qy          67 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 126
               |:|:|:|||:|||||||||| : |||||||||:||||:|| ||||| : ||||||| ||
Db         165 TAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITRIISKAKGQTRE 224

Qy         127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPPVLDSDGS 184
               ||||||||  :||::::||:|||| |||| || |||: |||  || ||:||||  | ||:
Db         225 PQVYTLPPHAEELSRSKVSITCLVIGFYPPDIDVEWQRNGQPEPEGNYRTTPPQQDVDGT 284

Qy         185 FFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
               :||||| :|||: || | :| |:||||||||||||:| :|||
Db         285 YFLYSKFSVDKASWQGGGIFQCAVMHEALHNHYTQKSISKTPGK 328


RESULT 13
I47162
Ig gamma 4 chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47162
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47162
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-277 <KAC>
A;Cross-references: UNIPARC:UPI0000115527; EMBL:U03782; NID:g433129; PIDN:AAA52220.1; P
C;Genetics:
A;Gene: IgG4
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;82-151/Domain: immunoglobulin homology <IMM>

  Query Match             67.4%;  Score 903.5;  DB 2;  Length 277;
  Best Local Similarity   71.1%;  Pred. No. 2.3e-62;
  Matches  167;  Conservative   30;  Mismatches   31;  Indels    7;  Gaps    4;

Qy           1 MDK---THTCPPCP-APELLG-GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 55
               :||   | | ||||  |   | ||| |:|||||||||||||:||||||||| |:|||:
Db          43 VDKRVGTKTKPPCPICPACEGPGPSAFIFPPKPKDTLMISRTPKVTCVVVDVSQENPEVQ 102

Qy          56 FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 115
               |:|||||||| |:|:|:|||:|||||||||| : |||||||||:||||:|| ||||| :
Db         103 FSWYVDGVEVHTAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITR 162

Qy         116 TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYK 173
               ||||||| |||||||||| :||::::|:||||| |||| || |||: ||| || ||:
Db         163 IISKAKGQTREPQVYTLPPPTEELSRSKVTLTCLVTGFYPPDIDVEWQRNGQPEPEGNYR 222

Qy         174 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
               |||| | ||::|||||| |||: ||:|: | |:|||||||||||||:  :|||
Db         223 TTPPQQDVDGTYFLYSKLAVDKASWQRGDTFQCAVMHEALHNHYTQKSIFKTPGK 277


RESULT 14
G2GP
Ig gamma-2 chain C region - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 09-Jul-2004
C;Accession: A94553; A90352; A90359; A90384; A90385; A02151
R;Trischmann, T.M.
submitted to the Atlas, April 1975
A;Reference number: A94553
```

A;Accession: A94553
A;Molecule type: protein
A;Residues: 1-3 <TRI>
A;Cross-references: UNIPROT:P01862; UNIPARC:UPI000017379E
R;Birshtein, B.K.; Hussain, Q.Z.; Cebra, J.J.
Biochemistry 10, 18-25, 1971
A;Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Am
A;Reference number: A90352; MUID:71058471; PMID:5538606
A;Accession: A90352
A;Molecule type: protein
A;Residues: 4-68 <BIR>
A;Cross-references: UNIPARC:UPI000017379F
R;Turner, K.J.; Cebra, J.J.
Biochemistry 10, 9-17, 1971
A;Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. Ami
A;Reference number: A90359; MUID:71058486; PMID:5538616
A;Accession: A90359
A;Molecule type: protein
A;Residues: 69-133;312-329 <TUR>
A;Cross-references: UNIPARC:UPI00001737A0; UNIPARC:UPI00001737A1
R;Tracey, D.E.; Cebra, J.J.
Biochemistry 13, 4796-4803, 1974
A;Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibodies.
A;Reference number: A90384; MUID:75036072; PMID:4429665
A;Accession: A90384
A;Molecule type: protein
A;Residues: 134-226 <TRA>
A;Cross-references: UNIPARC:UPI00001737A2
R;Trischmann, T.M.; Cebra, J.J.
Biochemistry 13, 4804-4811, 1974
A;Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibodies.
A;Reference number: A90385; MUID:75036073; PMID:4609467
A;Accession: A90385
A;Molecule type: protein
A;Residues: 227-311 <TR2>
A;Cross-references: UNIPARC:UPI00001737A3
R;Oliveira, B.; Lamm, M.E.
Biochemistry 10, 26-31, 1971
A;Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.
A;Reference number: A90354; MUID:71058474; PMID:4922544
A;Contents: annotation; disulfide bonds
A;Note: Cys-16 is involved in a heavy-light chain bond
A;Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds
C;Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;21-81/Domain: immunoglobulin homology <IM1>
F;135-204/Domain: immunoglobulin homology <IM2>
F;241-310/Domain: immunoglobulin homology <IM3>
F;28-79/Disulfide bonds: #status experimental
F;142-202/Disulfide bonds: #status experimental
F;178/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;248-308/Disulfide bonds: #status experimental

Query Match 66.3%; Score 889; DB 1; Length 329;
Best Local Similarity 72.3%; Pred. No. 3.8e-61;
Matches 162; Conservative 24; Mismatches 36; Indels 2; Gaps 1;

```
Qy          6 TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 65
              ||| || || |||||||:|||||||||||| || ||||||||| ::|||:| |:||   |
Db        106 TCPKCPPPENLGGPSVFIFPPKPKDTLMISLTPRVTCVVVDVSQDEPEVQFTWFVDNKPV 165

Qy         66 HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 125
               ||:|||| |||||:|:|| ||| : ||||| |||:|||| |||||||||||||| || ||
Db        166 GNAETKPRVEQYNTTFRVESVLPIQHQDWLRGKEFKCKVYNKALPAPIEKTISKTKGAPR 225

Qy        126 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP--ENNYKTTPPVLDSDG 183
               | |||||||||||:|::||:|||:  |:|:|| ||| ||  |  |  || |||: |:||
Db        226 MPDVYTLPPSRDELSKSKVSVTCLIINFFPADIHVEWASNRVPVSEKEYKNTPPIEDADG 285

Qy        184 SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG 227
```

```
                |:|||||||||||| | || |::|||||||||||| |||::| |||
Db        286 SYFLYSKLTVDKSAWDQGTVYTCSVMHEALHNHVTQKAISRSPG 329


RESULT 15
I47158
Ig gamma 1 chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47158
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47158
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115523; EMBL:U03778; NID:g433121; PIDN:AAA52216.1;
C;Genetics:
A;Gene: IgG1
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain: immunoglobulin homology <IMM>

Query Match 66.0%; Score 885.5; DB 2; Length 328;
Best Local Similarity 72.4%; Pred. No. 7e-61;
Matches 163; Conservative 27; Mismatches 32; Indels 3; Gaps 2;

Qy          6 TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 65
              ||| ||  | : |||||:|||||||||||||:|||||||||||| |  ||:|:|||||||
Db        105 TCPICPGCE-VAGPSVFIFPPKPKDTLMISQTPEVTCVVVDVSKEHAEVQFSWYVDGVEV 163

Qy         66 HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 125
               | |:|:|:|||:|||||||||| : ||||| |||:||||:|  |||||  :|||||  || |
Db        164 HTAETRPKEEQFNSTYRVVSVLPIQHQDWLKGKEFKCKVNNVDLPAPITRTISKAIGQSR 223

Qy        126 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPPVLDSDG 183
              |||||||||  :||::::|:|||||  |||| || |||:||||  |||  |:||||  | ||
Db        224 EPQVYTLPPPAEELSRSKVTLTCLVIGFYPPDIHVEWKSNGQPEPENTYRTTPPQQDVDG 283

Qy        184 SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              :|||||||  |||:||  |: | |:|||||||||||||:| : ||
Db        284 TFFLYSKLAVDKARWDHGDKFECAVMHEALHNHYTQKSISKTQGK 328


Search completed: April 4, 2006, 13:17:27
Job time : 41.4123 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 188.806 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-6
Perfect score: 1341
Sequence: 1 MDKTHTCPPCPAPELLGGPS.....KGGGGGIEGPTLRQWLAARA 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1233	91.9	330	1	IGHG1_HUMAN	P01857 homo sapien
2	1233	91.9	465	2	Q6GMX6_HUMAN	Q6gmx6 homo sapien
3	1233	91.9	466	2	Q6IN78_HUMAN	Q6in78 homo sapien
4	1233	91.9	469	2	Q569F4_HUMAN	Q569f4 homo sapien
5	1233	91.9	469	2	Q7Z7P5_HUMAN	Q7z7p5 homo sapien
6	1233	91.9	470	2	Q7Z5W1_HUMAN	Q7z5w1 homo sapien
7	1233	91.9	470	2	Q6PJA4_HUMAN	Q6pja4 homo sapien
8	1233	91.9	472	2	Q6N089_HUMAN	Q6n089 homo sapien
9	1233	91.9	475	2	Q5EFE5_HUMAN	Q5efe5 homo sapien
10	1233	91.9	475	2	Q6GMW7_HUMAN	Q6gmw7 homo sapien
11	1233	91.9	476	2	Q6GMX1_HUMAN	Q6gmx1 homo sapien
12	1233	91.9	679	2	Q96PQ8_HUMAN	Q96pq8 homo sapien
13	1229	91.6	473	2	Q6P055_HUMAN	Q6p055 homo sapien
14	1229	91.6	475	2	Q6MZQ6_HUMAN	Q6mzq6 homo sapien
15	1229	91.6	480	2	Q6N094_HUMAN	Q6n094 homo sapien
16	1229	91.6	481	2	Q6N097_HUMAN	Q6n097 homo sapien
17	1229	91.6	482	2	Q7Z351_HUMAN	Q7z351 homo sapien
18	1227	91.5	348	2	Q6PYX1_HUMAN	Q6pyx1 homo sapien
19	1227	91.5	473	2	Q6MZV7_HUMAN	Q6mzv7 homo sapien
20	1227	91.5	478	2	Q6PI81_HUMAN	Q6pi81 homo sapien
21	1227	91.5	480	2	Q6PJF1_HUMAN	Q6pjf1 homo sapien
22	1226	91.4	466	2	Q6N096_HUMAN	Q6n096 homo sapien
23	1222	91.1	475	2	Q6N095_HUMAN	Q6n095 homo sapien
24	1222	91.1	544	2	Q6PJ95_HUMAN	Q6pj95 homo sapien
25	1216	90.7	487	2	Q65ZL2_9MURI	Q65zl2 mus sp. fv/
26	1172	87.4	475	2	Q5RE17_PONPY	Q5re17 pongo pygma
27	1146	85.5	354	2	Q86TT2_HUMAN	Q86tt2 homo sapien
28	1146	85.5	518	2	Q6N030_HUMAN	Q6n030 homo sapien
29	1146	85.5	519	2	Q5EBM2_HUMAN	Q5ebm2 homo sapien
30	1142.5	85.2	326	1	IGHG2_HUMAN	P01859 homo sapien
31	1142.5	85.2	417	2	Q6N093_HUMAN	Q6n093 homo sapien
32	1142	85.2	521	2	Q8N4Y9_HUMAN	Q8n4y9 homo sapien
33	1139.5	85.0	464	2	Q6MZU6_HUMAN	Q6mzu6 homo sapien
34	1137.5	84.8	465	2	Q6P6C4_HUMAN	Q6p6c4 homo sapien
35	1135	84.6	327	1	IGHG4_HUMAN	P01861 homo sapien
36	1135	84.6	473	2	Q8TC63_HUMAN	Q8tc63 homo sapien
37	1131	84.3	509	2	Q8NF17_HUMAN	Q8nf17 homo sapien
38	1128.5	84.2	470	2	Q68CN4_HUMAN	Q68cn4 homo sapien
39	1126	84.0	290	1	IGHG3_HUMAN	P01860 homo sapien
40	1126	84.0	476	2	Q6MZX7_HUMAN	Q6mzx7 homo sapien
41	918.5	68.5	323	1	GC_RABIT	P01870 oryctolagus
42	909	67.8	337	2	Q95M34_HORSE	Q95m34 equus cabal
43	889	66.3	329	1	IGHG2_CAVPO	P01862 cavia porce
44	845.5	63.0	329	1	GC3_MOUSE	P22436 mus musculu
45	845.5	63.0	470	2	Q7TMK1_MOUSE	Q7tmk1 mus musculu

ALIGNMENTS

RESULT 1
IGHG1_HUMAN
ID IGHG1_HUMAN STANDARD; PRT; 330 AA.
AC P01857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ig gamma-1 chain C region.
GN Name=IGHG1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=82274238; PubMed=6287432;
RA Ellison J.W., Berson B.J., Hood L.E.;
RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene.";
RL Nucleic Acids Res. 10:4071-4079(1982).
RN [2]
RP PROTEIN SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
RX MEDLINE=71064024; PubMed=5489771;
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA Waxdal M.J., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino
RT acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";
RL Biochemistry 9:3161-3170(1970).
RN [3]
RP PROTEIN SEQUENCE OF 136-329 (EU).
RX MEDLINE=71064025; PubMed=5530842;
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,
RA Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino
RT acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";
RL Biochemistry 9:3171-3181(1970).
RN [4]
RP PROTEIN SEQUENCE (MYELOMA PROTEIN NIE).
RX MEDLINE=77070269; PubMed=826475;
RA Ponstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic
RT peptides of the H-chain, alignment of the tryptic peptides and
RT discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [5]
RP PROTEIN SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [6]
RP DISULFIDE BONDS.

```
RX   MEDLINE=71064027; PubMed=4923144;
RA   Gall W.E., Edelman G.M.;
RT   "The covalent structure of a human gamma G-immunoglobulin. X.
RT   Intrachain disulfide bonds.";
RL   Biochemistry 9:3188-3196(1970).
RN   [7]
RP   DISULFIDE BONDS.
RX   MEDLINE=77070267; PubMed=1002129;
RA   Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT   "Rule of antibody structure. The primary structure of a monoclonal
RT   IgG1 immunoglobulin (myeloma protein Nie), I: purification and
RT   characterization of the protein, the L- and H-chains, the cyanogen
RT   bromide cleavage products, and the disulfide bridges.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX   MEDLINE=81208100; PubMed=7236608;
RA   Deisenhofer J.;
RT   "Crystallographic refinement and atomic models of a human Fc fragment
RT   and its complex with fragment B of protein A from Staphylococcus
RT   aureus at 2.9- and 2.8-A resolution.";
RL   Biochemistry 20:2361-2370(1981).
CC   -!- MISCELLANEOUS: Nie has the G1M(17) allotypic marker, 97-K, and the
CC       G1M(1) markers, 239-D and 241-L. KOL and EU sequences have the
CC       G1M(3) marker and the G1M (non-1) markers.
CC   -!- MISCELLANEOUS: Nie also differs in the amidation states of 35,
CC       116, 198, 269 and 272.
CC   -!- MISCELLANEOUS: EU also differs in the amidation states of residues
CC       155, 166, 177, 195, 198, 269, and 272 and in the order of residues
CC       268-272.
CC   -!- MISCELLANEOUS: KOL also differs in the amidation states of
CC       residues 198, 267 and 272.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; J00228; AAC82527.1; ALT_INIT; Genomic_DNA.
DR   PIR; A93433; GHHU.
DR   PDB; 1AJ7; X-ray; H=1-103.
DR   PDB; 1AQK; X-ray; H=1-103.
DR   PDB; 1D5B; X-ray; B/H=1-101.
DR   PDB; 1D5I; X-ray; H=1-101.
DR   PDB; 1D6V; X-ray; H=1-101.
DR   PDB; 1DN2; X-ray; A/B=120-326.
DR   PDB; 1E4K; X-ray; A/B=106-330.
DR   PDB; 1FC1; X-ray; A/B=106-329.
DR   PDB; 1FC2; X-ray; D=106-329.
DR   PDB; 1FCC; X-ray; A=121-326.
DR   PDB; 1HZH; X-ray; H/K=1-330.
DR   PDB; 1I7Z; X-ray; B/D=1-103.
DR   PDB; 1IIS; X-ray; A/B=107-330.
DR   PDB; 1IIX; X-ray; A/B=107-330.
DR   PDB; 1L6X; X-ray; A=120-326.
DR   PDB; 1OQX; X-ray; A/B=119-330.
DR   PDB; 1T83; X-ray; A/B=107-330.
DR   PDB; 2RCS; X-ray; H=1-103.
DR   HGNC; HGNC:5525; IGHG1.
DR   MIM; 147100; -.
DR   GO; GO:0005624; C:membrane fraction; NAS.
DR   GO; GO:0003823; F:antigen binding; TAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF07654; C1-set; 3.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS00290; IG_MHC; 2.
KW   3D-structure; Direct protein sequencing; Glycoprotein;
KW   Immunoglobulin C region; Immunoglobulin domain.
FT   REGION        1     98       CH1.
FT   REGION       99    110       Hinge.
FT   REGION      111    223       CH2.
FT   REGION      224    330       CH3.
FT   CARBOHYD    180    180       N-linked (GlcNAc. . .).
FT   DISULFID     27     83
FT   DISULFID    103    103       Interchain (with light chain).
FT   DISULFID    109    109       Interchain (with heavy chain).
FT   DISULFID    112    112       Interchain (with heavy chain).
FT   DISULFID    144    204
FT   DISULFID    250    308
FT   VARIANT      97     97       K -> R (in G1M(3) marker).
FT                                /FTId=VAR_003886.
FT   VARIANT     239    239       D -> E (in G1M(non-1) marker).
FT                                /FTId=VAR_003887.
FT   VARIANT     241    241       L -> M (in G1M(non-1) marker).
FT                                /FTId=VAR_003888.
FT   NON_TER       1      1
FT   STRAND       23     24
FT   STRAND       26     33
FT   STRAND       38     38
FT   STRAND       41     41
FT   TURN         42     45
FT   TURN         48     49
FT   STRAND       50     52
FT   STRAND       57     58
FT   TURN         59     61
FT   STRAND       62     71
FT   HELIX        73     75
FT   TURN         76     78
FT   STRAND       82     87
FT   TURN         88     91
FT   STRAND       92     97
FT   TURN        102    103
FT   STRAND      122    126
FT   HELIX       130    134
FT   TURN        136    137
FT   STRAND      141    149
FT   STRAND      157    162
FT   TURN        163    164
FT   STRAND      165    167
FT   STRAND      171    172
FT   STRAND      176    177
FT   TURN        179    180
FT   STRAND      183    190
FT   HELIX       193    197
FT   TURN        198    199
FT   STRAND      202    207
FT   TURN        209    210
FT   STRAND      215    219
FT   STRAND      227    227
FT   STRAND      230    234
FT   HELIX       238    242
FT   STRAND      245    256
FT   STRAND      261    266
FT   TURN        267    268
FT   STRAND      269    270
FT   STRAND      274    276
FT   STRAND      280    281
FT   TURN        283    284
FT   STRAND      287    296
FT   HELIX       297    301
FT   TURN        302    303
FT   STRAND      306    311
FT   TURN        313    314
FT   HELIX       316    318
FT   STRAND      319    324
SQ   SEQUENCE   330 AA;  36106 MW;  3770EE106C2FA33D CRC64;

  Query Match             91.9%;  Score 1233;  DB 1;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 223

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 283

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
                ||||||||||||||||||||||||||||||||||||||||||||||
Db        284 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 330
```

RESULT 2
Q6GMX6_HUMAN
ID Q6GMX6_HUMAN PRELIMINARY; PRT; 465 AA.
AC Q6GMX6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Primary B-Cells;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Primary B-Cells;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC073766; AAH73766.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 465 AA; 51083 MW; B3A9B7D0FDB1386E CRC64;

```
  Query Match             91.9%;  Score 1233;  DB 2;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 298

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 358

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 418

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
                ||||||||||||||||||||||||||||||||||||||||||||||
Db        419 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 465
```

RESULT 3
Q6IN78_HUMAN
ID Q6IN78_HUMAN PRELIMINARY; PRT; 466 AA.
AC Q6IN78;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE IGHG1 protein.
GN Name=IGHG1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RG NIH MGC Project;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC072419; AAH72419.1; -; mRNA.
DR HSSP; P01861; 1ADQ.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 3.

```
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   466 AA;  50854 MW;  53EB0BCEDE81076E CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 466;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 299

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 359

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 419

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        420 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 466


RESULT 4
Q569F4_HUMAN
ID   Q569F4_HUMAN PRELIMINARY;      PRT;   469 AA.
AC   Q569F4;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lymph;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lymph;
RG   NIH MGC Project;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC092518; AAH92518.1; -; mRNA.
SQ   SEQUENCE   469 AA;  51254 MW;  AC13448E3047784F CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 302

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 362

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 422

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        423 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 469


RESULT 5
Q7Z7P5_HUMAN
ID   Q7Z7P5_HUMAN PRELIMINARY;      PRT;   469 AA.
AC   Q7Z7P5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RG   NIH MGC Project;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC051328; AAH51328.1; -; mRNA.
DR   HSSP; P01857; 1HZH.
DR   SMR; Q7Z7P5; 20-469.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
```

```
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Immunoglobulin domain.
SQ   SEQUENCE   469 AA;  51395 MW;  C8D5BE12BAAF795C CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 302

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 362

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 422

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        423 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 469


RESULT 6
Q7Z5W1_HUMAN
ID   Q7Z5W1_HUMAN PRELIMINARY;      PRT;   470 AA.
AC   Q7Z5W1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RA   Strausberg R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC053984; AAH53984.1; -; mRNA.
DR   HSSP; P01857; 1HZH.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   470 AA;  51204 MW;  778CF34521483E1A CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 303

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 363

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 423

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        424 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 470


RESULT 7
Q6PJA4_HUMAN
ID   Q6PJA4_HUMAN PRELIMINARY;      PRT;   470 AA.
AC   Q6PJA4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RG   NIH MGC Project;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; BC018747; AAH18747.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   SMR; Q6PJA4; 20-470.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   470 AA;  51716 MW;  7B49556A11FD7D99 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 303

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 363

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 423

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        424 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 470


RESULT 8
Q6N089_HUMAN
ID   Q6N089_HUMAN PRELIMINARY;      PRT;   472 AA.
AC   Q6N089;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686P15220.
GN   Name=DKFZp686P15220;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Rectum tumor;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640627; CAE45781.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   472 AA;  51724 MW;  26CB340D0046D279 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 305

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 365

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 425

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        426 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 472


RESULT 9
Q5EFE5_HUMAN
ID   Q5EFE5_HUMAN PRELIMINARY;      PRT;   475 AA.
AC   Q5EFE5;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Anti-RhD monoclonal T125 gamma1 heavy chain precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Gaucher C., Klein P., Beliard R.;
RT   "Sequence determination of the recombinant human anti-RhD monoclonal
RT   antibody T125.";
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY894992; AAW82028.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   Pfam; PF07686; V-set; 1.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Signal.
FT   SIGNAL       1     19       Potential.
FT   CHAIN       20    475       anti-RhD monoclonal T125 gamma1 heavy
FT                               chain.
SQ   SEQUENCE   475 AA;  52362 MW;  1367D400DC7D2859 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 475;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
```

```
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 475
```

RESULT 10
Q6GMW7_HUMAN
ID Q6GMW7_HUMAN PRELIMINARY; PRT; 475 AA.
AC Q6GMW7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC073782; AAH73782.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 475 AA; 51987 MW; 2A1FE55D736860F8 CRC64;

Query Match 91.9%; Score 1233; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 4.4e-91;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 475
```

RESULT 11
Q6GMX1_HUMAN
ID Q6GMX1_HUMAN PRELIMINARY; PRT; 476 AA.
AC Q6GMX1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC073773; AAH73773.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF07654; C1-set; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGv; 1.
DR PROSITE; PS50835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 476 AA; 52286 MW; 622AABA5C62DDE9D CRC64;

Query Match 91.9%; Score 1233; DB 2; Length 476;

```
  Best Local Similarity   100.0%;  Pred. No. 4.4e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 309

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 369

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 429

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 476


RESULT 12
Q96PQ8_HUMAN
ID   Q96PQ8_HUMAN PRELIMINARY;      PRT;   679 AA.
AC   Q96PQ8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VII active site mutant immunoconjugate.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21477448; PubMed=11593034; DOI=10.1073/pnas.201420298;
RA   Hu Z., Garen A.;
RT   "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT   cells for immunotherapy in mouse models of prostatic cancer.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Hu Z., Garen A.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF272774; AAK58686.2; -; mRNA.
DR   HSSP; P08709; 1KLI.
DR   SMR; Q96PQ8; 39-180, 191-444, 447-679.
DR   Ensembl; ENSG00000057593; Homo sapiens.
DR   GO; GO:0005576; C:extracellular region; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR001254; Peptidase_S1_S6.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF07654; C1-set; 2.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00594; Gla; 1.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00407; IGc1; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLA_1; UNKNOWN_1.
DR   PROSITE; PS50998; GLA_2; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
SQ   SEQUENCE   679 AA;  75552 MW;  0B0023AE70A067A1 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 679;
  Best Local Similarity   100.0%;  Pred. No. 7e-91;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 512

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 572

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 632

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 679


RESULT 13
Q6P055_HUMAN
ID   Q6P055_HUMAN PRELIMINARY;      PRT;   473 AA.
AC   Q6P055;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
```

```
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RA   Strausberg R.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC065820; AAH65820.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   473 AA;  51344 MW;  9816D56A77129B57 CRC64;

  Query Match             91.6%;  Score 1229;  DB 2;  Length 473;
  Best Local Similarity   99.6%;  Pred. No. 9.2e-91;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 306

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 366

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 426

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||| ||||||||||||||||
Db        427 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 473


RESULT 14
Q6MZQ6_HUMAN
ID   Q6MZQ6_HUMAN PRELIMINARY;      PRT;   475 AA.
AC   Q6MZQ6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686G11190.
GN   Name=DKFZp686G11190;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Esophagus tumor;
RG   The German cDNA Consortium;
RA   Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA   Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640947; CAE45972.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   SMR; Q6MZQ6; 20-475.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   475 AA;  52043 MW;  B7EAE255A26F4B8E CRC64;

  Query Match             91.6%;  Score 1229;  DB 2;  Length 475;
  Best Local Similarity   99.6%;  Pred. No. 9.2e-91;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||| ||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 475


RESULT 15
Q6N094_HUMAN
ID   Q6N094_HUMAN PRELIMINARY;      PRT;   480 AA.
AC   Q6N094;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686O01196.
GN   Name=DKFZp686O01196;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Esophagus tumor;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640622; CAE45776.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   480 AA;  52612 MW;  225247F3D35AEC18 CRC64;

  Query Match             91.6%;  Score 1229;  DB 2;  Length 480;
  Best Local Similarity   99.6%;  Pred. No. 9.4e-91;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        254 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 313

Qy         62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 373

Qy        122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 433

Qy        182 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
              |||||||||||||||||||||||||||| ||||||||||||||||
Db        434 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 480
```

Search completed: April 4, 2006, 13:15:18
Job time : 188.806 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 122.53 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-12
Perfect score: 1341
Sequence: 1 MIEGPTLRQWLAARAGGGGG.....MHEALHNHYTQKSLSLSPGK 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1341	100.0	247	3	AAB16961	Aab16961 TMP-Fc pr
2	1341	100.0	247	5	ABB73414	Abb73414 TMP-Fc am
3	1336	99.6	269	3	AAB16960	Aab16960 TMP-TMP-F
4	1336	99.6	269	5	ABB73413	Abb73413 TMP-TMP-F
5	1270	94.7	275	9	ADW97940	Adw97940 Human TWE
6	1270	94.7	322	9	ADW97944	Adw97944 TWEAKR:Gl
7	1270	94.7	339	9	ADW97943	Adw97943 TWEAKR:1K
8	1269	94.6	248	3	AAB17952	Aab17952 TNF-alpha
9	1269	94.6	248	5	ABB73420	Abb73420 TNF-alpha
10	1269	94.6	252	6	ABJ38339	Abj38339 TALL-1 in
11	1268.5	94.6	252	6	ABJ38336	Abj38336 TALL-1 in
12	1268.5	94.6	293	6	ABJ38344	Abj38344 TALL-1 in
13	1268.5	94.6	293	8	ADQ76789	Adq76789 AGP3 pept
14	1267	94.5	252	3	AAB17956	Aab17956 VEGF anta
15	1267	94.5	252	5	ABB73424	Abb73424 VEGF anta
16	1267	94.5	283	5	AAE15488	Aae15488 Human BCM
17	1266.5	94.4	252	6	ABJ38343	Abj38343 TALL-1 in
18	1266.5	94.4	293	6	ABJ38345	Abj38345 TALL-1 in
19	1264.5	94.3	248	6	ABJ38332	Abj38332 TALL-1 in
20	1264	94.3	248	3	AAB17954	Aab17954 IL-1 anta
21	1264	94.3	248	5	ABB73422	Abb73422 Interleuk
22	1264	94.3	397	5	AAE15498	Aae15498 Human TAC
23	1263	94.2	244	7	ADN59685	Adn59685 Vector 20
24	1263	94.2	250	3	AAB17958	Aab17958 MMP inhib
25	1263	94.2	250	5	ABB73426	Abb73426 MMP inhib
26	1263	94.2	253	3	AAB16965	Aab16965 EMP-Fc pr
27	1263	94.2	253	5	ABB73416	Abb73416 EPO mimet
28	1263	94.2	255	9	AEA18571	Aea18571 Amino aci
29	1263	94.2	277	3	AAB16966	Aab16966 EMP-EMP-F
30	1263	94.2	281	5	AAE15489	Aae15489 Mouse BCM
31	1263	94.2	282	9	ADW97969	Adw97969 Human TWE
32	1263	94.2	462	7	ADC98598	Adc98598 Human ang
33	1263	94.2	489	7	ADC98596	Adc98596 Human ang
34	1263	94.2	588	7	ADC98594	Adc98594 Human ang
35	1263	94.2	648	7	ADC98590	Adc98590 Human ang
36	1263	94.2	665	7	ADC98592	Adc98592 Human ang
37	1263	94.2	697	7	ADC98614	Adc98614 Human ang
38	1263	94.2	705	7	ADC98588	Adc98588 Human ang
39	1263	94.2	726	7	ADC98586	Adc98586 Human ang
40	1263	94.2	883	7	ADC98568	Adc98568 Human ang
41	1260.5	94.0	252	6	ABJ38338	Abj38338 TALL-1 in
42	1259.5	93.9	248	6	ABJ38333	Abj38333 TALL-1 in
43	1258.5	93.8	252	6	ABJ38341	Abj38341 TALL-1 in
44	1258	93.8	358	9	AEB46697	Aeb46697 Human FSH
45	1258	93.8	377	9	AEB46699	Aeb46699 Human FSH

ALIGNMENTS

RESULT 1
AAB16961
ID AAB16961 standard; protein; 247 AA.
XX
AC AAB16961;
XX
DT 31-OCT-2000 (first entry)
XX
DE TMP-Fc protein sequence SEQ ID NO:12.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69447.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 21; Page 188-189; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,

CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 247 AA;

Query Match 100.0%; Score 1341; DB 3; Length 247;
Best Local Similarity 100.0%; Pred. No. 5.6e-93;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIEGPTLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIEGPTLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT 60

Qy         61 CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 120

Qy        121 CKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 180

Qy        181 WESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 WESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 240

Qy        241 LSLSPGK 247
              |||||||
Db        241 LSLSPGK 247
```

RESULT 2
ABB73414
ID ABB73414 standard; protein; 247 AA.
XX
AC ABB73414;
XX
DT 05-APR-2002 (first entry)
XX
DE TMP-Fc amino acid SEQ ID NO:12.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35764.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 21; Fig 10; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 247 AA;

Query Match 100.0%; Score 1341; DB 5; Length 247;
Best Local Similarity 100.0%; Pred. No. 5.6e-93;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MIEGPTLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MIEGPTLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT 60

Qy         61 CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 120

Qy        121 CKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 180

Qy        181 WESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 WESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 240

Qy        241 LSLSPGK 247
              |||||||
Db        241 LSLSPGK 247
```

RESULT 3
AAB16960
ID AAB16960 standard; protein; 269 AA.
XX
AC AAB16960;
XX
DT 31-OCT-2000 (first entry)
XX

```
DE   TMP-TMP-Fc protein sequence SEQ ID NO:10.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW   inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW   thrombosis; pharmaceutical.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-US025044.
XX
PR   23-OCT-1998;   98US-0105371P.
PR   22-OCT-1999;   99US-00428082.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham J,  Boone TC;
XX
DR   WPI; 2000-350702/30.
DR   N-PSDB; AAA69446.
XX
PT   Novel composition of matter comprising an Fc domain and pharmacologically
PT   active peptides, useful for treating cancer and autoimmune diseases.
XX
PS   Example 2; Page 185-186; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC   (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC   P3, and P4 = are each independently sequences of pharmacologically active
CC   peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC   c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC   of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC   thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC   cells from the present invention can be used for producing pharmaceutical
CC   compositions. The compositions are useful for treating cancer, asthma,
CC   thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC   a Fab domain) can provide a longer half-life or incorporate functions
CC   such as Fc receptor binding, protein A binding, complement fixation, and
CC   possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC   AAB18003 represent nucleotide and amino acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 269 AA;

  Query Match             99.6%;  Score 1336;  DB 3;  Length 269;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-92;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IEGPTLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTC 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 IEGPTLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTC 83

Qy         62 VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC 143

Qy        122 KVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 KVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 203

Qy        182 ESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        204 ESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL 263

Qy        242 SLSPGK 247
              ||||||
Db        264 SLSPGK 269


RESULT 4
ABB73413
ID   ABB73413 standard; protein; 269 AA.
XX
AC   ABB73413;
XX
DT   05-APR-2002  (first entry)
XX
DE   TMP-TMP-Fc amino acid SEQ ID NO:10.
XX
KW   Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW   erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW   TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW   TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW   MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW   cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW   antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW   neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW   cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW   sleep disorder; neurological degenerative disease; anaemia;
KW   thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW   Fanconi's syndrome.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200183525-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014310.
XX
PR   03-MAY-2000; 2000US-00563286.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham JC,  Boone TC,  Gudas JM;
XX
DR   WPI; 2002-130313/17.
DR   N-PSDB; ABL35763.
XX
PT   Novel vehicle-peptide molecule or its multimers useful for treating
PT   inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT   diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS   Example 2; Fig 9; 176pp; English.
XX
CC   The present invention describes a vehicle-peptide molecule (I) or its
CC   multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC   cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC   antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC   neuroprotective activities. (I) can be used as a therapeutic or
CC   prophylactic agent as well as for screening purposes. (I) is useful for
CC   diagnosing diseases characterised by dysfunction of their associated
CC   protein of interest, for identifying normal or abnormal proteins of
CC   interest, as a part of diagnostic kit to detect the presence of their
CC   proteins of interest in a biological sample. Additionally, (I) is useful
CC   for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC   rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC   infertility, and neurological degenerative diseases. (I), comprising EPO-
CC   mimetic compounds are useful for treating disorders characterised by low
CC   red blood cell levels such as anaemia. The TPO-mimetic comprising
CC   compounds are useful for treating conditions that involve an existing
CC   megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC   deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC   tumour which result in thrombocytopaenia, systemic lupus erythematosus,
```

```
CC   and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC   represent amino acid and nucleic acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 269 AA;

  Query Match             99.6%;  Score 1336;  DB 5;  Length 269;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-92;
  Matches  246;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 IEGPTLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTC 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 IEGPTLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTC 83

Qy         62 VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 VVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC 143

Qy        122 KVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 KVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 203

Qy        182 ESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 ESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL 263

Qy        242 SLSPGK 247
              ||||||
Db        264 SLSPGK 269


RESULT 5
ADW97940
ID   ADW97940 standard; protein; 275 AA.
XX
AC   ADW97940;
XX
DT   21-APR-2005  (first entry)
XX
DE   Human TWEAKR - Gly5- IgG1 Fc portion fusion protein, TWEAKR:Gly5:Fc.
XX
KW   TWEAK protein; TREPA; Apo3L; TWEAK receptor; radiotherapy; chemotherapy;
KW   pharmaceutical; delivery mechanism; antagonist; angiogenesis inhibition;
KW   transgenic animal; transgenic plant; protein interaction;
KW   animal disease model; angiogenesis disorder; antiangiogenic; solid tumor;
KW   cytostatic; neoplasm; opthalmological; inflammation; antiinflammatory;
KW   fusion protein; immunoglobulin; igg; fc receptor.
XX
OS   Homo sapiens.
OS   Chimeric.
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Region          1
FT                   /note= "Transcription start site region (N-terminal
FT                   region)"
FT   Region          2. .43
FT                   /note= "Human TWEAK receptor"
FT   Region          44. .48
FT                   /note= "Pentaglycine linker"
FT   Region          49. .275
FT                   /note= "Human IgG1 Fc protein"
XX
PN   WO2005010045-A1.
XX
PD   03-FEB-2005.
XX
PF   23-JUL-2004; 2004WO-US023904.
XX
PR   24-JUL-2003; 2003US-0490036P.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Wiley SR;
XX
DR   WPI; 2005-123128/13.
XX
PT   New fusion proteins comprising multimeric soluble TWEAK receptor
PT   fragments and an oligomerization domain, useful for antagonizing TWEAK
PT   receptor or for treating diseases mediated by angiogenesis, e.g. solid
PT   tumors or inflammation.
XX
PS   Claim 40; SEQ ID NO 15; 140pp; English.
XX
CC   The present invention provides methods and compositions relating to
CC   fusion proteins comprising multimeric soluble fragments of the major
CC   functional TWEAK (also called TREPA and Apo3L) receptor (TWEAKR) and an
CC   oligomerization domain. The invention is useful for inhibiting
CC   angiogenesis and for treating diseases such as solid tumors, ocular
CC   neovascularization and inflammatory conditions. The TWEAK receptor
CC   proteins of the invention are also used in the production of transgenic
CC   animals and plants. The present sequence is human TWEAK receptor (TWEAKR)
CC   - Gly5 - IgG1 Fc portion fusion protein. This fusion protein comprises a
CC   N-terminal methionine residue, human TWEAKR (residues 29 through 70 of
CC   the SEQ ID NO: 7), five glycine residues (linker) and the Fc portion of
CC   human IgG1. This sequence is used to illustrate an ELISA-style assay
CC   useful for determining the binding properties of TWEAK binding molecules.
XX
SQ   Sequence 275 AA;

  Query Match             94.7%;  Score 1270;  DB 9;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-87;
  Matches  234;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         13 ARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE 72
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 AAAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE 100

Qy         73 VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 160

Qy        133 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK 220

Qy        193 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 275


RESULT 6
ADW97944
ID   ADW97944 standard; protein; 322 AA.
XX
AC   ADW97944;
XX
DT   21-APR-2005  (first entry)
XX
DE   TWEAKR:Gly5:TWEAKR:Gly5:Fc fusion protein.
XX
KW   TWEAK protein; TREPA; Apo3L; TWEAK receptor; radiotherapy; chemotherapy;
KW   pharmaceutical; delivery mechanism; antagonist; angiogenesis inhibition;
KW   transgenic animal; transgenic plant; protein interaction;
KW   animal disease model; angiogenesis disorder; antiangiogenic; solid tumor;
KW   cytostatic; neoplasm; opthalmological; inflammation; antiinflammatory;
KW   fusion protein; immunoglobulin; igg; fc receptor.
XX
OS   Homo sapiens.
OS   Chimeric.
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Region          1
FT                   /note= "Transcription start site region (N-terminal
FT                   region)"
```

```
FT   Region          2. .43
FT                   /note= "Human TWEAK receptor"
FT   Region          44. .48
FT                   /note= "Pentaglycine linker"
FT   Region          49. .90
FT                   /note= "Human TWEAK receptor"
FT   Region          91. .95
FT                   /note= "Pentaglycine linker"
FT   Region          96. .322
FT                   /note= "Human IgG1 Fc protein"
XX
PN   WO2005010045-A1.
XX
PD   03-FEB-2005.
XX
PF   23-JUL-2004; 2004WO-US023904.
XX
PR   24-JUL-2003; 2003US-0490036P.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Wiley SR;
XX
DR   WPI; 2005-123128/13.
XX
PT   New fusion proteins comprising multimeric soluble TWEAK receptor
PT   fragments and an oligomerization domain, useful for antagonizing TWEAK
PT   receptor or for treating diseases mediated by angiogenesis, e.g. solid
PT   tumors or inflammation.
XX
PS   Claim 40; SEQ ID NO 19; 140pp; English.
XX
CC   The present invention provides methods and compositions relating to
CC   fusion proteins comprising multimeric soluble fragments of the major
CC   functional TWEAK (also called TREPA and Apo3L) receptor (TWEAKR) and an
CC   oligomerization domain. The invention is useful for inhibiting
CC   angiogenesis and for treating diseases such as solid tumors, ocular
CC   neovascularization and inflammatory conditions. The TWEAK receptor
CC   proteins of the invention are also used in the production of transgenic
CC   animals and plants. The present sequence is human TWEAK receptor (TWEAKR)
CC   - Gly5 - TWEAKR - Gly5- IgG1 Fc portion fusion protein. This fusion
CC   protein comprises a N-terminal methionine residue, human TWEAKR,
CC   pentaglycine linker, human TWEAKR, pentaglycine linker and the Fc portion
CC   of human IgG1 where the TWEAK receptor corresponds to residues 29 through
CC   70 of the SEQ ID NO: 7. This sequence is used to illustrate an ELISA-
CC   style assay useful for determining the binding properties of TWEAK
CC   binding molecules.
XX
SQ   Sequence 322 AA;

  Query Match             94.7%;  Score 1270;  DB 9;  Length 322;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-87;
  Matches  234;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         13 ARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE 72
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 AAAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE 147

Qy         73 VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 207

Qy        133 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK 267

Qy        193 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 322


RESULT 7
ADW97943
ID   ADW97943 standard; protein; 339 AA.
XX
AC   ADW97943;
XX
DT   21-APR-2005  (first entry)
XX
DE   TWEAKR:1KPEG:TWEAKR:Gly5:Fc fusion protein.
XX
KW   TWEAK protein; TREPA; Apo3L; TWEAK receptor; radiotherapy; chemotherapy;
KW   pharmaceutical; delivery mechanism; antagonist; angiogenesis inhibition;
KW   transgenic animal; transgenic plant; protein interaction;
KW   animal disease model; angiogenesis disorder; antiangiogenic; solid tumor;
KW   cytostatic; neoplasm; opthalmological; inflammation; antiinflammatory;
KW   fusion protein; immunoglobulin; igg; fc receptor.
XX
OS   Homo sapiens.
OS   Chimeric.
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Region          1
FT                   /note= "Transcription start site region (N-terminal
FT                   region)"
FT   Region          2. .43
FT                   /note= "Human TWEAK receptor"
FT   Region          44. .65
FT                   /note= "Linker"
FT   Region          66. .107
FT                   /note= "Human TWEAK receptor"
FT   Region          108. .112
FT                   /note= "Pentaglycine linker"
FT   Region          113. .339
FT                   /note= "Human IgG1 Fc protein"
XX
PN   WO2005010045-A1.
XX
PD   03-FEB-2005.
XX
PF   23-JUL-2004; 2004WO-US023904.
XX
PR   24-JUL-2003; 2003US-0490036P.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Wiley SR;
XX
DR   WPI; 2005-123128/13.
XX
PT   New fusion proteins comprising multimeric soluble TWEAK receptor
PT   fragments and an oligomerization domain, useful for antagonizing TWEAK
PT   receptor or for treating diseases mediated by angiogenesis, e.g. solid
PT   tumors or inflammation.
XX
PS   Claim 40; SEQ ID NO 18; 140pp; English.
XX
CC   The present invention provides methods and compositions relating to
CC   fusion proteins comprising multimeric soluble fragments of the major
CC   functional TWEAK (also called TREPA and Apo3L) receptor (TWEAKR) and an
CC   oligomerization domain. The invention is useful for inhibiting
CC   angiogenesis and for treating diseases such as solid tumors, ocular
CC   neovascularization and inflammatory conditions. The TWEAK receptor
CC   proteins of the invention are also used in the production of transgenic
CC   animals and plants. The present sequence is human TWEAK receptor (TWEAKR)
CC   - 1KPEG - TWEAKR - Gly5- IgG1 Fc portion fusion protein. This fusion
CC   protein comprises a N-terminal methionine residue, human TWEAKR, linker,
CC   human TWEAKR, pentaglycine linker and the Fc portion of human IgG1 where
CC   the TWEAK receptor corresponds to residues 29 through 70 of the SEQ ID
CC   NO: 7. This sequence is used to illustrate an ELISA-style assay useful
CC   for determining the binding properties of TWEAK binding molecules.
XX
SQ   Sequence 339 AA;

  Query Match             94.7%;  Score 1270;  DB 9;  Length 339;
```

```
  Best Local Similarity   99.6%;  Pred. No. 1.8e-87;
  Matches  234;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         13 ARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE 72
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 AAAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE 164

Qy         73 VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 224

Qy        133 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK 192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYK 284

Qy        193 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 TTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 339


RESULT 8
AAB17952
ID   AAB17952 standard; protein; 248 AA.
XX
AC   AAB17952;
XX
DT   31-OCT-2000  (first entry)
XX
DE   TNF-alpha inhibitor-Fc fusion protein sequence SEQ ID NO:1058.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW   inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW   thrombosis; pharmaceutical.
XX
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-US025044.
XX
PR   23-OCT-1998;   98US-0105371P.
PR   22-OCT-1999;   99US-00428082.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham J,  Boone TC;
XX
DR   WPI; 2000-350702/30.
DR   N-PSDB; AAA69502.
XX
PT   Novel composition of matter comprising an Fc domain and pharmacologically
PT   active peptides, useful for treating cancer and autoimmune diseases.
XX
PS   Example 4; Page 571-572; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC   (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC   P3, and P4 = are each independently sequences of pharmacologically active
CC   peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC   c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC   of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC   thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC   cells from the present invention can be used for producing pharmaceutical
CC   compositions. The compositions are useful for treating cancer, asthma,
CC   thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC   a Fab domain) can provide a longer half-life or incorporate functions
CC   such as Fc receptor binding, protein A binding, complement fixation, and
CC   possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC   AAB18003 represent nucleotide and amino acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 248 AA;

  Query Match             94.6%;  Score 1269;  DB 3;  Length 248;
  Best Local Similarity   98.7%;  Pred. No. 1.5e-87;
  Matches  234;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         11 LAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHED 70
              | | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 LGHRPGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHED 71

Qy         71 PEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 PEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA 131

Qy        131 PIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN 191

Qy        191 YKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 YKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 248


RESULT 9
ABB73420
ID   ABB73420 standard; protein; 248 AA.
XX
AC   ABB73420;
XX
DT   05-APR-2002  (first entry)
XX
DE   TNF-alpha inhibitor-Fc fusion nucleic acid SEQ ID NO:1057.
XX
KW   Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW   erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW   TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW   TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW   MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW   cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW   antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW   neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW   cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW   sleep disorder; neurological degenerative disease; anaemia;
KW   thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW   Fanconi's syndrome.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200183525-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014310.
XX
PR   03-MAY-2000; 2000US-00563286.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham JC,  Boone TC,  Gudas JM;
XX
DR   WPI; 2002-130313/17.
DR   N-PSDB; ABL35770.
XX
PT   Novel vehicle-peptide molecule or its multimers useful for treating
PT   inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
```

PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Example 4; Fig 20A-B; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 248 AA;

Query Match 94.6%; Score 1269; DB 5; Length 248;
Best Local Similarity 98.7%; Pred. No. 1.5e-87;
Matches 234; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         11 LAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHED 70
              |  | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 LGHRPGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHED 71

Qy         71 PEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA 130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 PEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA 131

Qy        131 PIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN 190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN 191

Qy        191 YKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 YKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 248
```

RESULT 10
ABJ38339
ID ABJ38339 standard; protein; 252 AA.
XX
AC ABJ38339;
XX
DT 12-JUN-2003 (first entry)
XX
DE TALL-1 inhibitory protein SEQ ID No 118.
XX
KW TALL-1-binding protein; TALL-1; B-cell-mediated autoimmune disease;
KW systemic lupus erythematosus; B-cell-mediated cancer; lymphoma;
KW inflammation; rheumatoid arthritis; acute pancreatitis; atherosclerosis;
KW Alzheimer's disease; asthma; cachexia; cirrhosis; diabetes; osteoporosis;
KW glomerulonephritis; Hashimoto's thyroiditis; ischaemic injury; psoriasis;
KW multiple myeloma; multiple sclerosis; Parkinson's disease; vasculitis;
KW gene therapy.
XX
OS Unidentified.
XX
PN WO200292620-A2.
XX
PD 21-NOV-2002.
XX
PF 13-MAY-2002; 2002WO-US015273.
XX
PR 11-MAY-2001; 2001US-0290196P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Min H, Hsu H;
XX
DR WPI; 2003-156719/15.
XX
PT New TALL-1-binding polypeptide, useful for modulating the activity of
PT TALL-1 and in treating, preventing or diagnosing a B-cell-mediated
PT autoimmune diseases, cancers or lymphomas.
XX
PS Example 2; Page 68; 236pp; English.
XX
CC The invention relates to a novel TALL-1-binding polypeptide comprising a
CC defined sequence in the specification. The composition is useful in
CC modulating the activity of TALL-1, and in treating, preventing,
CC ameliorating, diagnosing or prognosing a B-cell-mediated autoimmune
CC disease (e.g. systemic lupus erythematosus) or B-cell-mediated cancer or
CC lymphoma. The composition may also be used in treating inflammations
CC (e.g. rheumatoid arthritis), acute pancreatitis, Alzheimer's disease,
CC asthma, atherosclerosis, cachexia, cirrhosis, diabetes,
CC glomerulonephritis, Hashimoto's thyroiditis, ischaemic injury, multiple
CC myeloma, multiple sclerosis, osteoporosis, Parkinson's disease, psoriasis
CC and vasculitis. Disorders may be treated with the novel composition using
CC gene therapy. This sequence represents a TALL-1 inhibitory protein of the
CC invention
XX
SQ Sequence 252 AA;

Query Match 94.6%; Score 1269; DB 6; Length 252;
Best Local Similarity 96.3%; Pred. No. 1.5e-87;
Matches 233; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```
Qy          6 TLRQWLAARAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 65
              | ::|     |||| ||||||||||||||||||||||||||||||||||||||||||
Db         11 TYKEWCQFNGGGGGVDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 70

Qy         66 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 130

Qy        126 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 190

Qy        186 QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP 250

Qy        246 GK 247
              ||
Db        251 GK 252
```

RESULT 11
ABJ38336
ID ABJ38336 standard; protein; 252 AA.
XX
AC ABJ38336;
XX
DT 12-JUN-2003 (first entry)
XX
DE TALL-1 inhibitory protein SEQ ID No 115.
XX
KW TALL-1-binding protein; TALL-1; B-cell-mediated autoimmune disease;
KW systemic lupus erythematosus; B-cell-mediated cancer; lymphoma;
KW inflammation; rheumatoid arthritis; acute pancreatitis; atherosclerosis;
KW Alzheimer's disease; asthma; cachexia; cirrhosis; diabetes; osteoporosis;
KW glomerulonephritis; Hashimoto's thyroiditis; ischaemic injury; psoriasis;

KW multiple myeloma; multiple sclerosis; Parkinson's disease; vasculitis;
KW gene therapy.
XX
OS Unidentified.
XX
PN WO200292620-A2.
XX
PD 21-NOV-2002.
XX
PF 13-MAY-2002; 2002WO-US015273.
XX
PR 11-MAY-2001; 2001US-0290196P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Min H, Hsu H;
XX
DR WPI; 2003-156719/15.
XX
PT New TALL-1-binding polypeptide, useful for modulating the activity of
PT TALL-1 and in treating, preventing or diagnosing a B-cell-mediated
PT autoimmune diseases, cancers or lymphomas.
XX
PS Example 2; Page 67; 236pp; English.
XX
CC The invention relates to a novel TALL-1-binding polypeptide comprising a
CC defined sequence in the specification. The composition is useful in
CC modulating the activity of TALL-1, and in treating, preventing,
CC ameliorating, diagnosing or prognosing a B-cell-mediated autoimmune
CC disease (e.g. systemic lupus erythematosus) or B-cell-mediated cancer or
CC lymphoma. The composition may also be used in treating inflammations
CC (e.g. rheumatoid arthritis), acute pancreatitis, Alzheimer's disease,
CC asthma, atherosclerosis, cachexia, cirrhosis, diabetes,
CC glomerulonephritis, Hashimoto's thyroiditis, ischaemic injury, multiple
CC myeloma, multiple sclerosis, osteoporosis, Parkinson's disease, psoriasis
CC and vasculitis. Disorders may be treated with the novel composition using
CC gene therapy. This sequence represents a TALL-1 inhibitory protein of the
CC invention
XX
SQ Sequence 252 AA;

Query Match 94.6%; Score 1268.5; DB 6; Length 252;
Best Local Similarity 96.7%; Pred. No. 1.7e-87;
Matches 234; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

```
Qy          7 LRQWLAARAGGGGG-DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 65
              ::||:    ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         11 IKQWVCDPLGGGGGVDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 70

Qy         66 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         71 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 130

Qy        126 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 190

Qy        186 QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP 250

Qy        246 GK 247
              ||
Db        251 GK 252
```

RESULT 12
ABJ38344
ID ABJ38344 standard; protein; 293 AA.
XX
AC ABJ38344;
XX
DT 12-JUN-2003 (first entry)
XX
DE TALL-1 inhibitory protein SEQ ID No 123.
XX
KW TALL-1-binding protein; TALL-1; B-cell-mediated autoimmune disease;
KW systemic lupus erythematosus; B-cell-mediated cancer; lymphoma;
KW inflammation; rheumatoid arthritis; acute pancreatitis; atherosclerosis;
KW Alzheimer's disease; asthma; cachexia; cirrhosis; diabetes; osteoporosis;
KW glomerulonephritis; Hashimoto's thyroiditis; ischaemic injury; psoriasis;
KW multiple myeloma; multiple sclerosis; Parkinson's disease; vasculitis;
KW gene therapy.
XX
OS Unidentified.
XX
PN WO200292620-A2.
XX
PD 21-NOV-2002.
XX
PF 13-MAY-2002; 2002WO-US015273.
XX
PR 11-MAY-2001; 2001US-0290196P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Min H, Hsu H;
XX
DR WPI; 2003-156719/15.
XX
PT New TALL-1-binding polypeptide, useful for modulating the activity of
PT TALL-1 and in treating, preventing or diagnosing a B-cell-mediated
PT autoimmune diseases, cancers or lymphomas.
XX
PS Claim 42; Page 68; 236pp; English.
XX
CC The invention relates to a novel TALL-1-binding polypeptide comprising a
CC defined sequence in the specification. The composition is useful in
CC modulating the activity of TALL-1, and in treating, preventing,
CC ameliorating, diagnosing or prognosing a B-cell-mediated autoimmune
CC disease (e.g. systemic lupus erythematosus) or B-cell-mediated cancer or
CC lymphoma. The composition may also be used in treating inflammations
CC (e.g. rheumatoid arthritis), acute pancreatitis, Alzheimer's disease,
CC asthma, atherosclerosis, cachexia, cirrhosis, diabetes,
CC glomerulonephritis, Hashimoto's thyroiditis, ischaemic injury, multiple
CC myeloma, multiple sclerosis, osteoporosis, Parkinson's disease, psoriasis
CC and vasculitis. Disorders may be treated with the novel composition using
CC gene therapy. This sequence represents a TALL-1 inhibitory protein of the
CC invention
XX
SQ Sequence 293 AA;

Query Match 94.6%; Score 1268.5; DB 6; Length 293;
Best Local Similarity 96.7%; Pred. No. 2e-87;
Matches 234; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

```
Qy          7 LRQWLAARAGGGGG-DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 65
              ::||:    ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         52 IKQWVCDPLGGGGGVDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 111

Qy         66 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 171

Qy        126 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 231

Qy        186 QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP 291

Qy        246 GK 247
              ||
Db        292 GK 293
```

```
RESULT 13
ADQ76789
ID   ADQ76789 standard; protein; 293 AA.
XX
AC   ADQ76789;
XX
DT   07-OCT-2004  (first entry)
XX
DE   AGP3 peptibody, useful in immune and inflammatory disease therapy.
XX
KW   AGP3; peptibody; antiinflammatory; antiarthritic; antipsoriatic;
KW   immunosuppressive; dermatological; combination therapy;
KW   co-stimulatory factor.
XX
OS   Unidentified.
XX
PN   WO2004060911-A2.
XX
PD   22-JUL-2004.
XX
PF   24-DEC-2003; 2003WO-US041378.
XX
PR   30-DEC-2002; 2002US-0437405P.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Khare SD;
XX
DR   WPI; 2004-534365/51.
XX
PT   Treating an inflammatory, or autoimmune condition, e.g interleukin-1 or
PT   tumor necrosis factor-alpha comprises administering an IL-1 or TNF-alpha
PT   inhibitor and at least one of a B7 inhibitor and a CD28 inhibitor.
XX
PS   Claim 69; SEQ ID NO 1; 203pp; English.
XX
CC   The present sequence is the protein sequence of AGP3 peptibody, an AGP3
CC   tandem dimer peptide-Fc fusion. A claimed method for treating an
CC   inflammatory or an autoimmune condition comprises administering at least
CC   one of an AGP3 inhibitor, a BAFFR inhibitor and a TACI inhibitor
CC   comprising the present sequence, and at least one of a B7 inhibitor and a
CC   CD28 inhibitor. Alternatively, the AGP3 inhibitor, BAFFR inhibitor or
CC   TACI inhibitor is used in combination with an interleukin-1 inhibitor
CC   such as KIN2 ADQ76791 or a tumour necrosis factor-alpha inhibitor such as
CC   sTNFR-1 ADQ76792. The AGP3 inhibitor prevents binding of AGP3 to BAFFR
CC   and TACI. The inflammatory or autoimmune condition is selected from
CC   rheumatoid arthritis, psoriatic arthritis, systemic lupus erythematosus,
CC   graft rejection, psoriasis and inflammatory bowel disease. In examples
CC   from the invention, the combination of KIN2 or sTNFR-1 and AGP3 peptibody
CC   reduced collagen-induced arthritis in susceptible B10.RIII mice for an
CC   extended period compared to KIN2 alone.
XX
SQ   Sequence 293 AA;

  Query Match             94.6%;  Score 1268.5;  DB 8;  Length 293;
  Best Local Similarity   96.7%;  Pred. No. 2e-87;
  Matches  234;  Conservative    3;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          7 LRQWLAARAGGGGG-DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 65
              ::||:    ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db         52 IKQWVCDPLGGGGGVDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 111

Qy         66 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 VSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN 171

Qy        126 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 231

Qy        186 QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP 291

Qy        246 GK 247
              ||
Db        292 GK 293


RESULT 14
AAB17956
ID   AAB17956 standard; protein; 252 AA.
XX
AC   AAB17956;
XX
DT   31-OCT-2000  (first entry)
XX
DE   VEGF antagonist-Fc fusion protein sequence SEQ ID NO:1066.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW   inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW   thrombosis; pharmaceutical.
XX
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-US025044.
XX
PR   23-OCT-1998;   98US-0105371P.
PR   22-OCT-1999;   99US-00428082.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham J,  Boone TC;
XX
DR   WPI; 2000-350702/30.
DR   N-PSDB; AAA69506.
XX
PT   Novel composition of matter comprising an Fc domain and pharmacologically
PT   active peptides, useful for treating cancer and autoimmune diseases.
XX
PS   Example 6; Page 582-583; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC   (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC   P3, and P4 = are each independently sequences of pharmacologically active
CC   peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC   c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC   of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC   thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC   cells from the present invention can be used for producing pharmaceutical
CC   compositions. The compositions are useful for treating cancer, asthma,
CC   thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC   a Fab domain) can provide a longer half-life or incorporate functions
CC   such as Fc receptor binding, protein A binding, complement fixation, and
CC   possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC   AAB18003 represent nucleotide and amino acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 252 AA;

  Query Match             94.5%;  Score 1267;  DB 3;  Length 252;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-87;
  Matches  233;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy         14 RAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 73
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 RLGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 78

Qy         74 KFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 KFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE 138

Qy        134 KTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 KTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 198

Qy        194 TPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 252
```

```
RESULT 15
ABB73424
ID   ABB73424 standard; protein; 252 AA.
XX
AC   ABB73424;
XX
DT   05-APR-2002  (first entry)
XX
DE   VEGF antagonist-Fc fusion nucleic acid SEQ ID NO:1065.
XX
KW   Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW   erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW   TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW   TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW   MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW   cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW   antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW   neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW   cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW   sleep disorder; neurological degenerative disease; anaemia;
KW   thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW   Fanconi's syndrome.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200183525-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014310.
XX
PR   03-MAY-2000; 2000US-00563286.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham JC,  Boone TC,  Gudas JM;
XX
DR   WPI; 2002-130313/17.
DR   N-PSDB; ABL35774.
XX
PT   Novel vehicle-peptide molecule or its multimers useful for treating
PT   inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT   diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS   Example 6; Fig 24A-B; 176pp; English.
XX
CC   The present invention describes a vehicle-peptide molecule (I) or its
CC   multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC   cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC   antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC   neuroprotective activities. (I) can be used as a therapeutic or
CC   prophylactic agent as well as for screening purposes. (I) is useful for
CC   diagnosing diseases characterised by dysfunction of their associated
CC   protein of interest, for identifying normal or abnormal proteins of
CC   interest, as a part of diagnostic kit to detect the presence of their
CC   proteins of interest in a biological sample. Additionally, (I) is useful
CC   for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC   rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC   infertility, and neurological degenerative diseases. (I), comprising EPO-
CC   mimetic compounds are useful for treating disorders characterised by low
CC   red blood cell levels such as anaemia. The TPO-mimetic comprising
CC   compounds are useful for treating conditions that involve an existing
CC   megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC   deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC   tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC   and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC   represent amino acid and nucleic acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 252 AA;
```

```
  Query Match             94.5%;  Score 1267;  DB 5;  Length 252;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-87;
  Matches  233;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         14 RAGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 73
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 RLGGGGGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 78

Qy         74 KFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 KFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE 138

Qy        134 KTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 KTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 198

Qy        194 TPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 252
```

```
Search completed: April  4, 2006, 13:07:38
Job time : 123.53 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 40.4123 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-12
Perfect score: 1341
Sequence: 1 MIEGPTLRQWLAARAGGGGG.....MHEALHNHYTQKSLSLSPGK 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1233	91.9	255	4	S31866	Ig gamma-1 chain C
2	1233	91.9	330	1	GHHU	Ig gamma-1 chain C
3	1227	91.5	374	2	S69339	Ig heavy chain V r
4	1180	88.0	234	2	PT0207	Ig gamma chain C r
5	1146	85.5	377	2	A23511	Ig gamma-3 chain C
6	1144	85.3	377	2	A60764	Ig gamma-3 chain C
7	1142.5	85.2	326	1	G2HU	Ig gamma-2 chain C
8	1135.5	84.7	327	1	G4HU	Ig gamma-4 chain C
9	1121	83.6	289	1	G3HUWI	Ig gamma-3 heavy c
10	917.5	68.4	323	1	GHRB	Ig gamma chain C r
11	909	67.8	328	2	I47160	Ig gamma 2b chain
12	909	67.8	328	2	I47159	Ig gamma 2a chain
13	903	67.3	277	2	I47162	Ig gamma 4 chain c
14	889	66.3	329	1	G2GP	Ig gamma-2 chain C
15	886.5	66.1	328	2	I47158	Ig gamma 1 chain c
16	881	65.7	328	2	I47161	Ig gamma 3 chain c
17	855.5	63.8	470	2	S22080	Ig heavy chain pre
18	854.5	63.7	333	2	PS0018	Ig gamma-2b chain
19	846	63.1	308	2	C30554	Ig heavy chain C r
20	846	63.1	472	2	S31459	Ig gamma-1 chain -
21	845.5	63.0	329	1	G3MSC	Ig gamma-3 chain C
22	834.5	62.2	398	1	G3MSM	Ig gamma-3 chain C
23	827.5	61.7	444	2	PC4436	monoclonal antibod
24	824.5	61.5	326	2	PS0017	Ig gamma-1 chain C
25	817.5	61.0	324	1	G1MS	Ig gamma-1 chain C
26	812.5	60.6	393	1	G1MSM	Ig gamma-1 chain C
27	812	60.6	330	1	G2MSA	Ig gamma-2a chain
28	812	60.6	469	2	S37483	Ig gamma-2a chain
29	809.5	60.4	329	2	S00847	Ig gamma-2c chain
30	807	60.2	399	1	G2MSAM	Ig gamma-2a chain
31	802	59.8	335	1	G2MSAB	Ig gamma-2a chain
32	797	59.4	446	2	S40295	Ig gamma-2a chain
33	785.5	58.6	322	2	PS0019	Ig gamma-2a chain
34	779	58.1	474	1	G2MS11	Ig gamma-2b chain
35	774	57.7	405	1	G2MSBM	Ig gamma-2b chain
36	765.5	57.1	327	2	S06611	Ig gamma-2 chain C
37	757	56.5	475	2	S01321	Ig gamma-2b chain
38	707	52.7	180	2	I46732	Ig gamma heavy cha
39	577.5	43.1	249	2	S69340	Ig heavy chain VHI
40	574.5	42.8	218	2	A36040	Ig heavy chain V-I
41	571	42.6	152	2	S14236	Ig gamma-1 chain C
42	401.5	29.9	572	2	B46529	Ig Y heavy chain (
43	362	27.0	388	1	EHMS	Ig epsilon chain C
44	362	27.0	426	2	I36948	Ig epsilon-chain -
45	359	26.8	548	2	S38864	Ig epsilon chain C

ALIGNMENTS

RESULT 1
S31866
Ig gamma-1 chain C region - synthetic
C;Species: synthetic
A;Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C;Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C;Accession: S31866
R;Filpula, D.
submitted to the EMBL Data Library, February 1993
A;Description: Screeing method for protein-protein interactions of cloned gene product
A;Reference number: S31866
A;Accession: S31866
A;Molecule type: mRNA
A;Residues: 1-255 <FIL>
A;Cross-references: UNIPARC:UPI000011F41F; EMBL:X70421; NID:g33068; PIDN:CAA49866.1; P
C;Keywords: immunoglobulin
F;1-22/Region: Escherichia coli outer membrane protein A precursor
F;23-255/Region: human Ig gamma-1 chain C region

Query Match 91.9%; Score 1233; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.3e-88;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 88

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 148

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 208

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        209 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 255
```

RESULT 2
GHHU
Ig gamma-1 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 09-Jul-2004
C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146
R;Ellison, J.W.; Berson, B.J.; Hood, L.E.
Nucleic Acids Res. 10, 4071-4079, 1982
A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.
A;Reference number: A93433; MUID:82274238; PMID:6287432
A;Accession: A93433
A;Molecule type: DNA

A;Residues: 1-330 <ELL>
A;Cross-references: UNIPROT:P01857; UNIPARC:UPI0000034C0E; EMBL:Z17370
A;Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,
A;Note: Lys-330 is removed after translation
R;Harris, L.J.
submitted to the EMBL Data Library, October 1992
A;Reference number: S33904
A;Accession: S36861
A;Molecule type: DNA
A;Residues: 2-330 <HAR>
A;Cross-references: UNIPARC:UPI000013C6FE; EMBL:Z17370
R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of a
A;Reference number: S33887; MUID:83001943; PMID:6811139
A;Accession: S33887
A;Molecule type: DNA
A;Residues: 88-113;235-330 <TAK>
A;Cross-references: UNIPARC:UPI000017378B; UNIPARC:UPI000017378C; EMBL:Z17370
R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,
Biochemistry 9, 3161-3170, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen
A;Reference number: A90563; MUID:71064024; PMID:5489771
A;Contents: myeloma protein Eu
A;Accession: B90563
A;Molecule type: protein
A;Residues: 1-96,'R',98-135 <CUN>
A;Cross-references: UNIPARC:UPI000017378D
A;Note: this sequence has the G1m(3) marker, 97-Arg
R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid seque
A;Reference number: A90564; MUID:71064025; PMID:5530842
A;Contents: Eu
A;Accession: A90564
A;Molecule type: protein
A;Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',240,
A;Cross-references: UNIPARC:UPI000017378E
A;Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met
R;Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie),
igen Primaerstruktur.
A;Reference number: A91668; MUID:77070269; PMID:826475
A;Contents: myeloma protein Nie
A;Accession: B91668
A;Molecule type: protein
A;Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27
A;Cross-references: UNIPARC:UPI000017378F
A;Note: this sequence has the G1m(17) and G1m(1) markers
R;Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A;Title: Die Primaerstruktur des kristallisierbaren monoklonalen Immunglobulins IgG1 KOL
A;Reference number: A91723; MUID:83289131; PMID:6884994
A;Contents: myeloma protein KOL; disulfide bonds
A;Accession: A91723
A;Molecule type: protein
A;Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH
A;Cross-references: UNIPARC:UPI0000173790
A;Note: this sequence has the G1m(3) and G1m(non-1) markers
R;Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A;Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
A;Reference number: A90565; MUID:71064027; PMID:4923144
A;Contents: annotation; disulfide bonds
R;Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A;Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A;Reference number: A91667; MUID:77070267; PMID:1002129
A;Contents: annotation; disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG1
A;Cross-references: GDB:120085; OMIM:147100
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 114/1; 224/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into l
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;137-206/Domain: immunoglobulin homology <IM2>
F;243-310/Domain: immunoglobulin homology <IM3>
F;27-83,144-204,250-308/Disulfide bonds: #status experimental
F;103/Disulfide bonds: interchain (to light chain) #status experimental
F;109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F;180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 91.9%; Score 1233; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 7.4e-88;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       104 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163

Qy        81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       164 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 223

Qy       141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 283

Qy       201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db       284 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 330
```

RESULT 3
S69339
Ig heavy chain V region precursor - human
C;Species: Homo sapiens (man)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C;Accession: S69339; S72664
R;Khamlichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A;Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A;Reference number: S69339; MUID:95262687; PMID:7744049
A;Accession: S69339
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-374 <KHA>
A;Cross-references: UNIPARC:UPI0000176F24; EMBL:X81695
R;Khamlichi, A.A.
submitted to the EMBL Data Library, September 1994
A;Reference number: S72664
A;Accession: S72664
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-140,'C',142-374 <KH2>
A;Cross-references: UNIPARC:UPI0000176F25; EMBL:X81695
C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 91.5%; Score 1227; DB 2; Length 374;
Best Local Similarity 99.1%; Pred. No. 2.5e-87;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy        21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       148 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 207

Qy        81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       208 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 267
```

```
Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||:|:||||||||||||||||||||||||||||||||||||||||||
Db        268 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 327

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        328 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 374
```

RESULT 4
PT0207
Ig gamma chain C region - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C;Accession: PT0207
R;Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A;Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A;Reference number: PT0207; MUID:91287716; PMID:2062315
A;Accession: PT0207
A;Molecule type: mRNA
A;Residues: 1-234 <EHR>
A;Cross-references: UNIPARC:UPI0000176F05
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;48-117/Domain: immunoglobulin homology <IMM>

  Query Match 88.0%; Score 1180; DB 2; Length 234;
  Best Local Similarity 98.6%; Pred. No. 5.9e-84;
  Matches 217; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              | ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 DTTHTCPPCAAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 74

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 134

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        135 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYKTTPPVLDS 194

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 240
              |||||||||||||||||||||||||||||||||||||||||
Db        195 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKS 234
```

RESULT 5
A23511
Ig gamma-3 chain C region (allotype G3m(b)) - human
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
C;Accession: A23511
R;Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
Nucleic Acids Res. 14, 1779-1789, 1986
A;Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: co
A;Reference number: A23511; MUID:86148507; PMID:3081877
A;Accession: A23511
A;Molecule type: DNA
A;Residues: 1-377 <HUC>
A;Cross-references: UNIPARC:UPI000004718F; GB:X03604; GB:M12958; NID:g33070; PIDN:CAA272
C;Genetics:
A;Gene: GDB:IGHG3
A;Cross-references: GDB:119339; OMIM:147120
A;Map position: 14q32.33-14q32.33
A;Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: immunoglobulin
F;20-85/Domain: immunoglobulin homology <IMM>

  Query Match 85.5%; Score 1146; DB 2; Length 377;
  Best Local Similarity 92.5%; Pred. No. 4.6e-81;
  Matches 210; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

```
Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              |    || ||||||||||||||||||||||||||||||||||||||||||:| ||||
Db        151 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 210

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||||| |
Db        211 GVEVHNAKTKPREEQYNSTFRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKTK 270

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||:|:||||||||||||||||||||||:||||||| ||||:|||
Db        271 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDS 330

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||:|||||||||||||| :|||||||||||
Db        331 DGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPGK 377
```

RESULT 6
A60764
Ig gamma-3 chain C region, form LAT - human
C;Species: Homo sapiens (man)
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 31-Dec-2004
C;Accession: A60764
R;Huck, S.; Lefranc, G.; Lefranc, M.P.
Immunogenetics 30, 250-257, 1989
A;Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conve
A;Reference number: A60764; MUID:90007613; PMID:2571587
A;Accession: A60764
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-377 <HUC>
A;Cross-references: UNIPROT:Q8N4Y9; UNIPARC:UPI0000176F0B
C;Superfamily: immunoglobulin homology
C;Keywords: immunoglobulin
F;20-85/Domain: immunoglobulin homology <IMM>

  Query Match 85.3%; Score 1144; DB 2; Length 377;
  Best Local Similarity 92.5%; Pred. No. 6.5e-81;
  Matches 210; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

```
Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              |    || ||||||||||||||||||||||||||||||||||||||||||:| ||||
Db        151 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 210

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||:||||||||||||||||||||||||||||||||||||| |
Db        211 GVEVHNAKTKPREEQYNSTFRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKTK 270

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||:|:||||||||||||||||||||:||||||| ||||||||
Db        271 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPVLDS 330

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              ||||||||:|||||||||:|||||||||||||| :|||||||||||
Db        331 DGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNRFTQKSLSLSPGK 377
```

RESULT 7
G2HU
Ig gamma-2 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1981 #sequence_revision 13-Jun-1983 #text_change 09-Jul-2004
C;Accession: A93906; A92809; A90752; A93132; A02148
R;Ellison, J.; Hood, L.
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982
A;Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain c
A;Reference number: A93906; MUID:82197621; PMID:6804948
A;Accession: A93906
A;Molecule type: DNA
A;Residues: 1-326 <ELL>
A;Cross-references: UNIPROT:P01859; UNIPARC:UPI000003BFCC; GB:V00554; GB:J00230; NID:g1

```
A;Note: Lys-326 is probably removed posttranslationally
R;Wang, A.C.; Tung, E.; Fudenberg, H.H.
J. Immunol. 125, 1048-1054, 1980
A;Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f
A;Reference number: A92809; MUID:81007873; PMID:6774012
A;Contents: myeloma protein Til
A;Accession: A92809
A;Molecule type: protein
A;Residues: 1-19,'Q',21-57,'Z',59,'A',61-193,'D',195-325 <WAN>
A;Cross-references: UNIPARC:UPI0000173791
A;Note: Trp-156 is at or near the complement-binding site
R;Connell, G.E.; Parr, D.M.; Hofmann, T.
Can. J. Biochem. 57, 758-767, 1979
A;Title: The amino acid sequences of the three heavy chain constant region domains of a
A;Reference number: A90752; MUID:80001357; PMID:113060
A;Contents: myeloma protein Zie
A;Accession: A90752
A;Molecule type: protein
A;Residues: 1-24,'E',26-57,'EV',60-85;132-171,'ZZZ',175,'B',177-193,'D',195-196,'Q',198-
A;Cross-references: UNIPARC:UPI0000173792; UNIPARC:UPI0000173793
A;Note: this sequence has since been revised
R;Hofmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A;Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g
A;Reference number: A93132; MUID:80114419; PMID:118920
A;Contents: Zie
A;Accession: A93132
A;Molecule type: protein
A;Residues: 238-275 <HOF>
A;Cross-references: UNIPARC:UPI0000173794
R;Hofmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A;Reference number: A94591
A;Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268
A;Note: the revised sequence differs from that shown in having 60-Ala and in the amidati
ned
R;Milstein, C.; Frangione, B.
Biochem. J. 121, 217-225, 1971
A;Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.
A;Reference number: A90253; MUID:72033500; PMID:4940472
A;Contents: annotation; myeloma protein Sa, disulfide bonds
R;Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A;Title: Structural studies of immunoglobulin G.
A;Reference number: A93157; MUID:69064124; PMID:5782707
A;Contents: annotation; Sa, disulfide bonds
C;Genetics:
A;Gene: GDB:IGHG2
A;Cross-references: GDB:119338; OMIM:147110
A;Map position: 14q32.33-14q32.33
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;133-202/Domain: immunoglobulin homology <IM2>
F;239-306/Domain: immunoglobulin homology <IM3>
F;14/Disulfide bonds: interchain (to light chain) #status experimental
F;27-83,140-200,246-304/Disulfide bonds: #status experimental
F;102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F;176/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             85.2%;  Score 1142.5;  DB 1;  Length 326;
  Best Local Similarity   94.1%;  Pred. No. 7.1e-81;
  Matches  209;  Conservative    8;  Mismatches    4;  Indels    1;  Gaps    1;

Qy         26 CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 85
              |||||||  : |||||||||||||||||||||||||||||||||||:||||||||||
Db        106 CPPCPAPP-VAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVH 164

Qy         86 NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 145
              ||||||||||:|||:||||||||:|||||||||||||||| |||||||||| ||||||
Db        165 NAKTKPREEQFNSTFRVVSVLTVVHQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPRE 224

Qy        146 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 205
              |||||||||||:|:|||||||||||||||||||||||||||||||||||:||||||||
Db        225 PQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPMLDSDGSFF 284

Qy        206 LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||
Db        285 LYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 326


RESULT 8
G4HU
Ig gamma-4 chain C region - human
C;Species: Homo sapiens (man)
C;Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 09-Jul-2004
C;Accession: A90933; A90249; A02150
R;Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A;Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A;Reference number: A90933; MUID:83157104; PMID:6299662
A;Accession: A90933
A;Molecule type: DNA
A;Residues: 1-327 <ELL>
A;Cross-references: UNIPROT:P01861; UNIPARC:UPI0000047190
A;Note: the sequence was determined from the germline gene
R;Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A;Title: Human immunoglobulin sublclasses. Partial amino acid sequence of the constant
A;Reference number: A90249; MUID:70207560; PMID:4192699
A;Accession: A90249
A;Molecule type: protein
A;Residues: 1-30;81-326 <PIN>
A;Cross-references: UNIPARC:UPI0000173795; UNIPARC:UPI0000173796
C;Genetics:
A;Gene: GDB:IGHG4
A;Cross-references: GDB:119340; OMIM:147130
A;Map position: 14q32.33-14q32.33
A;Introns: 99/1; 111/1; 221/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into l
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-85/Domain: immunoglobulin homology <IM1>
F;99-110/Region: hinge
F;134-203/Domain: immunoglobulin homology <IM2>
F;240-307/Domain: immunoglobulin homology <IM3>
F;14/Disulfide bonds: interchain (to light chain) #status experimental
F;27-83,141-201,247-305/Disulfide bonds: #status predicted
F;106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F;177/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             84.7%;  Score 1135.5;  DB 1;  Length 327;
  Best Local Similarity   84.5%;  Pred. No. 2.5e-80;
  Matches  212;  Conservative    9;  Mismatches    7;  Indels   23;  Gaps    1;

Qy         20 GDKTHT----------------------CPPCPAPELLGGPSVFLFPPKPKDTLMISRT 56
              | ||:|                      || ||||| ||||||||||||||||||||
Db         77 GTKTYTCNVDHKPSNTKVDKRVESKYGPPCPSCPAPEFLGGPSVFLFPPKPKDTLMISRT 136

Qy         57 PEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG 116
              |||||||||| |||||:|||||||||||||||||:|||||||||||||||||||
Db        137 PEVTCVVVDVSQEDPEVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNG 196

Qy        117 KEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSD 176
              |||||||||| ||: |||||||||||||||||||::|:||||||||||||||||
Db        197 KEYKCKVSNKGLPSSIEKTISKAKGQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSD 256

Qy        177 IAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHY 236
              |||||||||||||||||||||||||||||||||:|||||||:|||||||||||||||||
Db        257 IAVEWESNGQPENNYKTTPPVLDSDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHY 316

Qy        237 TQKSLSLSPGK 247
              |||||||| ||
Db        317 TQKSLSLSPGK 327
```

```
Db        317 TQKSLSLSLGK 327


RESULT 9
G3HUWI
Ig gamma-3 heavy chain disease proteins - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C;Accession: A90442; A92219; A90198; A93915; A02149
R;Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A;Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-
A;Reference number: A90442; MUID:81021548; PMID:6774747
A;Contents: heavy chain disease protein Wis
A;Accession: A90442
A;Molecule type: protein
A;Residues: 1-289 <FRA>
A;Cross-references: UNIPARC:UPI0000173797
A;Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
A;Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
A;Note: the sequence of residues 42-76 was taken from the reference that follows
R;Michaelsen, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A;Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
A;Reference number: A92219; MUID:77118561; PMID:402363
A;Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W
A;Accession: A92219
A;Molecule type: protein
A;Residues: 12-97 <MIC>
A;Cross-references: UNIPARC:UPI0000173798
A;Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
idue segment (12-28)
A;Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter
R;Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A;Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
A;Reference number: A90198; MUID:77021516; PMID:823945
A;Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues
A;Accession: A90198
A;Molecule type: protein
A;Residues: 59-125,'EB',128-226,228-289 <WOL>
A;Cross-references: UNIPARC:UPI0000173799
A;Note: this protein lacks most of the V region, all of the CH1 region, and part of the
R;Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A;Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
A;Reference number: A93915; MUID:82247835; PMID:6808505
A;Contents: heavy chain disease protein Omm
A;Accession: A93915
A;Molecule type: mRNA
A;Residues: 12-70;72-114;116-125,'E',127-133,'L',135-136,'E',138,'Y',140-154,'D',156-157
A;Cross-references: UNIPARC:UPI000017379A; UNIPARC:UPI000017379B; UNIPARC:UPI000017379C;
A;Note: a carboxyl-terminal Lys is removed posttranslationally
A;Note: this sequence may represent an allelic form or another gamma chain subclass
C;Comment: The heavy chain disease protein Wis is shown.
C;Genetics:
A;Gene: GDB:IGHG3
A;Cross-references: GDB:119339; OMIM:147120
A;Map position: 14q32.33-14q32.33
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
F;203-270/Domain: immunoglobulin homology <IMM>
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

  Query Match             83.6%;  Score 1121;  DB 1;  Length 289;
  Best Local Similarity   90.3%;  Pred. No. 2.8e-79;
  Matches  204;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              |    || |||||||||||||||||||||||||||||||||||||||||||||||:| ||||
Db         64 DTPPPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVD 123

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||:|||||||||||:|:|||:|||||||||||||:||:|||||||||||||||||||||| |
Db        124 GVQVHNAKTKPREQQFNSTFRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTISKTK 183

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              |||||||||||||||:|:|||||||||||||||||||||||||:||||||| ||||:|||
Db        184 GQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESSGQPENNYNTTPPMLDS 243

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG 246
              |||||||||||||||||||||:|||||||||||| :|||||||||
Db        244 DGSFFLYSKLTVDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPG 289


RESULT 10
GHRB
Ig gamma chain C region - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 09-Jul-2004
C;Accession: A91749; A90290; A93928; A90245; A94416; A02161
R;Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A;Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplo
A;Reference number: A91749; MUID:84030930; PMID:6313520
A;Accession: A91749
A;Molecule type: mRNA
A;Residues: 1-323 <BER>
A;Cross-references: UNIPROT:P01870; UNIPARC:UPI000012B37D
A;Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Th
R;Pratt, D.M.; Mole, L.E.
Biochem. J. 151, 337-349, 1975
A;Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglo
A;Reference number: A90290; MUID:76135469; PMID:1243651
A;Accession: A90290
A;Molecule type: protein
A;Residues: 1-47,'E',49-71,'PV',72-128 <PRA>
A;Cross-references: UNIPARC:UPI00001737AB
R;Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
A;Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy chai
A;Reference number: A93928; MUID:83299917; PMID:6193512
A;Accession: A93928
A;Molecule type: mRNA
A;Residues: 88-103,'M',105-143,'E',145-184,'A',186,'E',188-266 <MAR>
A;Cross-references: UNIPARC:UPI000016C5ED; GB:M16426; NID:g165111; PIDN:AAA31289.1; PID
A;Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic mark
R;Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
Biochem. J. 116, 249-259, 1970
A;Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin
A;Reference number: A90245; MUID:70110015; PMID:5461106
A;Accession: A90245
A;Molecule type: protein
A;Residues: 132-143,'E',145-161 <FRU>
A;Cross-references: UNIPARC:UPI00001737AC
R;Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiksell
A;Reference number: A94416
A;Accession: A94416
A;Molecule type: protein
A;Residues: 129-131;155-172,'D',174-184,'A',186,'E',188-200,'D',202-217,'E',219-232,'Q'
A;Cross-references: UNIPARC:UPI00001737AD; UNIPARC:UPI00001737AE
A;Note: this has the e15 allotypic marker, 185-Ala
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into l
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;20-82/Domain: immunoglobulin homology <IM1>
F;130-199/Domain: immunoglobulin homology <IM2>
F;236-303/Domain: immunoglobulin homology <IM3>
F;173/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             68.4%;  Score 917.5;  DB 1;  Length 323;
  Best Local Similarity   72.0%;  Pred. No. 1.7e-63;
  Matches  167;  Conservative   28;  Mismatches   32;  Indels    5;  Gaps    2;
```

```
Qy          21 DKT---HTC--PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 75
               |||    ||  | || ||||||||||:||||||||||||||||||||||| :||||:|
Db          92 DKTVAPSTCSKPTCPPPELLGGPSVFIFPPKPKDTLMISRTPEVTCVVVDVSQDDPEVQF 151

Qy          76 NWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 135
               ||::  :|  |:   ||:|:||| |||| | : ||||| |||:|||| ||||||||||
Db         152 TWYINNEQVRTARPPLREQQFNSTIRVVSTLPITHQDWLRGKEFKCKVHNKALPAPIEKT 211

Qy         136 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 195
               ||||:||| ||:|||: | |:||:   |||||:: |||||||:|||| ||: |:||||||
Db         212 ISKARGQPLEPKVYTMGPPREELSSRSVSLTCMINGFYPSDISVEWEKNGKAEDNYKTTP 271

Qy         196 PVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
               |||||||:|||:||:|  | ||:|:||:||||||||||||||:| ||||
Db         272 AVLDSDGSYFLYNKLSVPTSEWQRGDVFTCSVMHEALHNHYTQKSISRSPGK 323
```

RESULT 11
I47160
Ig gamma 2b chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47160
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47160
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115525; EMBL:U03780; NID:g433125; PIDN:AAA52218.1; PI
C;Genetics:
A;Gene: IgG2b
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain: immunoglobulin homology <IMM>

```
  Query Match             67.8%;  Score 909;  DB 2;  Length 328;
  Best Local Similarity   72.3%;  Pred. No. 7.6e-63;
  Matches  167;  Conservative   29;  Mismatches   31;  Indels    4;  Gaps    3;

Qy          20 GDKTH-TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY 78
               | ||   || ||| |   |||||:||||||||||||||:||||||||| |:|||:|:||
Db          99 GTKTKPPCPICPACE-SPGPSVFIFPPKPKDTLMISRTPQVTCVVVDVSQENPEVQFSWY 157

Qy          79 VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK 138
               ||||||| |:|:|:|||:|||||||||| : |||||||||:||||:|| ||||| : |||
Db         158 VDGVEVHTAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITRIISK 217

Qy         139 AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPP 196
               |||| |||||||||  :||::::||:|||| |||| || |||: |||  || ||:||||
Db         218 AKGQTREPQVYTLPPHAEELSRSKVSITCLVIGFYPPDIDVEWQRNGQPEPEGNYRTTPP 277

Qy         197 VLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
               | ||::||||| :|||: || | :| |:|||||||||||||:| :|||
Db         278 QQDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALHNHYTQKSISKTPGK 328
```

RESULT 12
I47159
Ig gamma 2a chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47159
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a s
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47159
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115524; EMBL:U03779; NID:g433123; PIDN:AAA52217.1; PI
C;Genetics:
A;Gene: IgG2a
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain: immunoglobulin homology <IMM>

```
  Query Match             67.8%;  Score 909;  DB 2;  Length 328;
  Best Local Similarity   72.3%;  Pred. No. 7.6e-63;
  Matches  167;  Conservative   29;  Mismatches   31;  Indels    4;  Gaps    3;

Qy          20 GDKTH-TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY 78
               | ||   || ||| |   |||||:||||||||||||||:||||||||| |:|||:|:||
Db          99 GTKTKPPCPICPACE-SPGPSVFIFPPKPKDTLMISRTPQVTCVVVDVSQENPEVQFSWY 157

Qy          79 VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK 138
               ||||||| |:|:|:|||:|||||||||| : |||||||||:||||:|| ||||| : |||
Db         158 VDGVEVHTAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITRIISK 217

Qy         139 AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPP 196
               |||| |||||||||  :||::::||:|||| |||| || |||: |||  || ||:||||
Db         218 AKGQTREPQVYTLPPHAEELSRSKVSITCLVIGFYPPDIDVEWQRNGQPEPEGNYRTTPP 277

Qy         197 VLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
               | ||::||||| :|||: || | :| |:|||||||||||||:| :|||
Db         278 QQDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALHNHYTQKSISKTPGK 328
```

RESULT 13
I47162
Ig gamma 4 chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47162
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47162
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-277 <KAC>
A;Cross-references: UNIPARC:UPI0000115527; EMBL:U03782; NID:g433129; PIDN:AAA52220.1; PI
C;Genetics:
A;Gene: IgG4
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;82-151/Domain: immunoglobulin homology <IMM>

```
  Query Match             67.3%;  Score 903;  DB 2;  Length 277;
  Best Local Similarity   72.1%;  Pred. No. 1.8e-62;
  Matches  165;  Conservative   29;  Mismatches   31;  Indels    4;  Gaps    3;

Qy          23 THTCPPCP-APELLG-GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
               | | ||||  |   | ||| |:|||||||||||||||:||||||||| |:|||:|:||||
Db          49 TKTKPPCPICPACEGPGPSAFIFPPKPKDTLMISRTPKVTCVVVDVSQENPEVQFSWYVD 108

Qy          81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
               ||||| |:|:|:|||:|||||||||| : |||||||||:||||:|| ||||| : |||||
Db         109 GVEVHTAQTRPKEEQFNSTYRVVSVLPIQHQDWLNGKEFKCKVNNKDLPAPITRIISKAK 168

Qy         141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPPVL 198
               || ||||||||||  :||::::|:||||| |||| || |||: |||  || ||:||||
Db         169 GQTREPQVYTLPPPTEELSRSKVTLTCLVTGFYPPDIDVEWQRNGQPEPEGNYRTTPPQQ 228

Qy         199 DSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
               | ||::||||||| |||: ||:|: | |:|||||||||||||:  :|||
Db         229 DVDGTYFLYSKLAVDKASWQRGDTFQCAVMHEALHNHYTQKSIFKTPGK 277
```

RESULT 14
G2GP
Ig gamma-2 chain C region - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 09-Jul-2004

```
C;Accession: A94553; A90352; A90359; A90384; A90385; A02151
R;Trischmann, T.M.
submitted to the Atlas, April 1975
A;Reference number: A94553
A;Accession: A94553
A;Molecule type: protein
A;Residues: 1-3 <TRI>
A;Cross-references: UNIPROT:P01862; UNIPARC:UPI000017379E
R;Birshtein, B.K.; Hussain, Q.Z.; Cebra, J.J.
Biochemistry 10, 18-25, 1971
A;Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Am
A;Reference number: A90352; MUID:71058471; PMID:5538606
A;Accession: A90352
A;Molecule type: protein
A;Residues: 4-68 <BIR>
A;Cross-references: UNIPARC:UPI000017379F
R;Turner, K.J.; Cebra, J.J.
Biochemistry 10, 9-17, 1971
A;Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. Ami
A;Reference number: A90359; MUID:71058486; PMID:5538616
A;Accession: A90359
A;Molecule type: protein
A;Residues: 69-133;312-329 <TUR>
A;Cross-references: UNIPARC:UPI00001737A0; UNIPARC:UPI00001737A1
R;Tracey, D.E.; Cebra, J.J.
Biochemistry 13, 4796-4803, 1974
A;Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibodies.
A;Reference number: A90384; MUID:75036072; PMID:4429665
A;Accession: A90384
A;Molecule type: protein
A;Residues: 134-226 <TRA>
A;Cross-references: UNIPARC:UPI00001737A2
R;Trischmann, T.M.; Cebra, J.J.
Biochemistry 13, 4804-4811, 1974
A;Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibodies.
A;Reference number: A90385; MUID:75036073; PMID:4609467
A;Accession: A90385
A;Molecule type: protein
A;Residues: 227-311 <TR2>
A;Cross-references: UNIPARC:UPI00001737A3
R;Oliveira, B.; Lamm, M.E.
Biochemistry 10, 26-31, 1971
A;Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.
A;Reference number: A90354; MUID:71058474; PMID:4922544
A;Contents: annotation; disulfide bonds
A;Note: Cys-16 is involved in a heavy-light chain bond
A;Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds
C;Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F;21-81/Domain: immunoglobulin homology <IM1>
F;135-204/Domain: immunoglobulin homology <IM2>
F;241-310/Domain: immunoglobulin homology <IM3>
F;28-79/Disulfide bonds: #status experimental
F;142-202/Disulfide bonds: #status experimental
F;178/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;248-308/Disulfide bonds: #status experimental

  Query Match             66.3%;  Score 889;  DB 1;  Length 329;
  Best Local Similarity   72.3%;  Pred. No. 2.7e-61;
  Matches  162;  Conservative   24;  Mismatches   36;  Indels    2;  Gaps    1;

Qy         25 TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 84
              ||| || || |||||||:||||||||||||| || |||||||||| ::|||:| |:||   |
Db        106 TCPKCPPPENLGGPSVFIFPPKPKDTLMISLTPRVTCVVVDVSQDEPEVQFTWFVDNKPV 165

Qy         85 HNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 144
              ||:|||| ||||:|:|| ||| : ||||| |||:|||| |||||||||||||| || ||
Db        166 GNAETKPRVEQYNTTFRVESVLPIQHQDWLRGKEFKCKVYNKALPAPIEKTISKTKGAPR 225

Qy        145 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP--ENNYKTTPPVLDSDG 202
                 |||||||||||:|::||:|||:  |:|:|| ||| ||  |  | || ||| |:|||
Db        226 MPDVYTLPPSRDELSKSKVSVTCLIINFFPADIHVEWASNRVPVSEKEYKNTPPIEDADG 285

Qy        203 SFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG 246
              |:||||||||||| | || |::|||||||||| |||::| |||
Db        286 SYFLYSKLTVDKSAWDQGTVYTCSVMHEALHNHVTQKAISRSPG 329


RESULT 15
I47158
Ig gamma 1 chain constant region - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C;Accession: I47158
R;Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A;Reference number: I47158; MUID:95015845; PMID:7930579
A;Accession: I47158
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-328 <KAC>
A;Cross-references: UNIPARC:UPI0000115523; EMBL:U03778; NID:g433121; PIDN:AAA52216.1;
C;Genetics:
A;Gene: IgG1
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F;133-202/Domain: immunoglobulin homology <IMM>

  Query Match             66.1%;  Score 886.5;  DB 2;  Length 328;
  Best Local Similarity   71.0%;  Pred. No. 4.2e-61;
  Matches  164;  Conservative   28;  Mismatches   36;  Indels    3;  Gaps    2;

Qy         19 GGDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY 78
              |  : ||| ||  | : |||||:|||||||||||:|||||||||||| |  ||:|:||
Db         99 GIHQPQTCPICPGCE-VAGPSVFIFPPKPKDTLMISQTPEVTCVVVDVSKEHAEVQFSWY 157

Qy         79 VDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK 138
              ||||||| |:|:|:|||:|||||||||| : ||||| |||:||||:|  ||||| :||||
Db        158 VDGVEVHTAETRPKEEQFNSTYRVVSVLPIQHQDWLKGKEFKCKVNNVDLPAPITRTISK 217

Qy        139 AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTPP 196
              | || |||||||||||  :||::::|:||||| |||| || |||:|||| |||  |:||||
Db        218 AIGQSREPQVYTLPPPAEELSRSKVTLTCLVIGFYPPDIHVEWKSNGQPEPENTYRTTPP 277

Qy        197 VLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
                | ||:||||||| |||:||  |: | |:|||||||||||||:| : ||
Db        278 QQDVDGTFFLYSKLAVDKARWDHGDKFECAVMHEALHNHYTQKSISKTQGK 328


Search completed: April  4, 2006, 13:17:23
Job time : 41.4123 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 188.806 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-12
Perfect score: 1341
Sequence: 1 MIEGPTLRQWLAARAGGGGG.....MHEALHNHYTQKSLSLSPGK 247

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1233	91.9	330	1	IGHG1_HUMAN	P01857 homo sapien
2	1233	91.9	465	2	Q6GMX6_HUMAN	Q6gmx6 homo sapien
3	1233	91.9	466	2	Q6IN78_HUMAN	Q6in78 homo sapien
4	1233	91.9	469	2	Q569F4_HUMAN	Q569f4 homo sapien
5	1233	91.9	469	2	Q7Z7P5_HUMAN	Q7z7p5 homo sapien
6	1233	91.9	470	2	Q7Z5W1_HUMAN	Q7z5w1 homo sapien
7	1233	91.9	470	2	Q6PJA4_HUMAN	Q6pja4 homo sapien
8	1233	91.9	472	2	Q6N089_HUMAN	Q6n089 homo sapien
9	1233	91.9	475	2	Q5EFE5_HUMAN	Q5efe5 homo sapien
10	1233	91.9	475	2	Q6GMW7_HUMAN	Q6gmw7 homo sapien
11	1233	91.9	476	2	Q6GMX1_HUMAN	Q6gmx1 homo sapien
12	1233	91.9	679	2	Q96PQ8_HUMAN	Q96pq8 homo sapien
13	1229	91.6	473	2	Q6P055_HUMAN	Q6p055 homo sapien
14	1229	91.6	475	2	Q6MZQ6_HUMAN	Q6mzq6 homo sapien
15	1229	91.6	480	2	Q6N094_HUMAN	Q6n094 homo sapien
16	1229	91.6	481	2	Q6N097_HUMAN	Q6n097 homo sapien
17	1229	91.6	482	2	Q7Z351_HUMAN	Q7z351 homo sapien
18	1227	91.5	348	2	Q6PYX1_HUMAN	Q6pyx1 homo sapien
19	1227	91.5	473	2	Q6MZV7_HUMAN	Q6mzv7 homo sapien
20	1227	91.5	478	2	Q6PI81_HUMAN	Q6pi81 homo sapien
21	1227	91.5	480	2	Q6PJF1_HUMAN	Q6pjf1 homo sapien
22	1226	91.4	466	2	Q6N096_HUMAN	Q6n096 homo sapien
23	1222	91.1	475	2	Q6N095_HUMAN	Q6n095 homo sapien
24	1222	91.1	544	2	Q6PJ95_HUMAN	Q6pj95 homo sapien
25	1218	90.8	487	2	Q65ZL2_9MURI	Q65zl2 mus sp. fv/
26	1172	87.4	475	2	Q5RE17_PONPY	Q5re17 pongo pygma
27	1146	85.5	354	2	Q86TT2_HUMAN	Q86tt2 homo sapien
28	1146	85.5	518	2	Q6N030_HUMAN	Q6n030 homo sapien
29	1146	85.5	519	2	Q5EBM2_HUMAN	Q5ebm2 homo sapien
30	1142.5	85.2	326	1	IGHG2_HUMAN	P01859 homo sapien
31	1142.5	85.2	417	2	Q6N093_HUMAN	Q6n093 homo sapien
32	1142	85.2	521	2	Q8N4Y9_HUMAN	Q8n4y9 homo sapien
33	1139.5	85.0	464	2	Q6MZU6_HUMAN	Q6mzu6 homo sapien
34	1137.5	84.8	465	2	Q6P6C4_HUMAN	Q6p6c4 homo sapien
35	1135.5	84.7	327	1	IGHG4_HUMAN	P01861 homo sapien
36	1135.5	84.7	473	2	Q8TC63_HUMAN	Q8tc63 homo sapien
37	1131	84.3	509	2	Q8NF17_HUMAN	Q8nf17 homo sapien
38	1128.5	84.2	470	2	Q68CN4_HUMAN	Q68cn4 homo sapien
39	1126.5	84.0	476	2	Q6MZX7_HUMAN	Q6mzx7 homo sapien
40	1126	84.0	290	1	IGHG3_HUMAN	P01860 homo sapien
41	917.5	68.4	323	1	GC_RABIT	P01870 oryctolagus
42	909	67.8	337	2	Q95M34_HORSE	Q95m34 equus cabal
43	889	66.3	329	1	IGHG2_CAVPO	P01862 cavia porce
44	854.5	63.7	333	1	GCB_RAT	P20761 rattus norv
45	854.5	63.7	469	2	Q5M839_RAT	Q5m839 rattus norv

ALIGNMENTS

RESULT 1
IGHG1_HUMAN
ID IGHG1_HUMAN STANDARD; PRT; 330 AA.
AC P01857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ig gamma-1 chain C region.
GN Name=IGHG1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=82274238; PubMed=6287432;
RA Ellison J.W., Berson B.J., Hood L.E.;
RT "The nucleotide sequence of a human immunoglobulin C gamma1 gene.";
RL Nucleic Acids Res. 10:4071-4079(1982).
RN [2]
RP PROTEIN SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
RX MEDLINE=71064024; PubMed=5489771;
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA Waxdal M.J., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino
RT acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";
RL Biochemistry 9:3161-3170(1970).
RN [3]
RP PROTEIN SEQUENCE OF 136-329 (EU).
RX MEDLINE=71064025; PubMed=5530842;
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,
RA Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino
RT acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";
RL Biochemistry 9:3171-3181(1970).
RN [4]
RP PROTEIN SEQUENCE (MYELOMA PROTEIN NIE).
RX MEDLINE=77070269; PubMed=826475;
RA Ponstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic
RT peptides of the H-chain, alignment of the tryptic peptides and
RT discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [5]
RP PROTEIN SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [6]
RP DISULFIDE BONDS.

```
RX   MEDLINE=71064027; PubMed=4923144;
RA   Gall W.E., Edelman G.M.;
RT   "The covalent structure of a human gamma G-immunoglobulin. X.
RT   Intrachain disulfide bonds.";
RL   Biochemistry 9:3188-3196(1970).
RN   [7]
RP   DISULFIDE BONDS.
RX   MEDLINE=77070267; PubMed=1002129;
RA   Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT   "Rule of antibody structure. The primary structure of a monoclonal
RT   IgG1 immunoglobulin (myeloma protein Nie), I: purification and
RT   characterization of the protein, the L- and H-chains, the cyanogen
RT   bromide cleavage products, and the disulfide bridges.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX   MEDLINE=81208100; PubMed=7236608;
RA   Deisenhofer J.;
RT   "Crystallographic refinement and atomic models of a human Fc fragment
RT   and its complex with fragment B of protein A from Staphylococcus
RT   aureus at 2.9- and 2.8-A resolution.";
RL   Biochemistry 20:2361-2370(1981).
CC   -!- MISCELLANEOUS: Nie has the G1M(17) allotypic marker, 97-K, and the
CC       G1M(1) markers, 239-D and 241-L. KOL and EU sequences have the
CC       G1M(3) marker and the G1M (non-1) markers.
CC   -!- MISCELLANEOUS: Nie also differs in the amidation states of 35,
CC       116, 198, 269 and 272.
CC   -!- MISCELLANEOUS: EU also differs in the amidation states of residues
CC       155, 166, 177, 195, 198, 269, and 272 and in the order of residues
CC       268-272.
CC   -!- MISCELLANEOUS: KOL also differs in the amidation states of
CC       residues 198, 267 and 272.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; J00228; AAC82527.1; ALT_INIT; Genomic_DNA.
DR   PIR; A93433; GHHU.
DR   PDB; 1AJ7; X-ray; H=1-103.
DR   PDB; 1AQK; X-ray; H=1-103.
DR   PDB; 1D5B; X-ray; B/H=1-101.
DR   PDB; 1D5I; X-ray; H=1-101.
DR   PDB; 1D6V; X-ray; H=1-101.
DR   PDB; 1DN2; X-ray; A/B=120-326.
DR   PDB; 1E4K; X-ray; A/B=106-330.
DR   PDB; 1FC1; X-ray; A/B=106-329.
DR   PDB; 1FC2; X-ray; D=106-329.
DR   PDB; 1FCC; X-ray; A=121-326.
DR   PDB; 1HZH; X-ray; H/K=1-330.
DR   PDB; 1I7Z; X-ray; B/D=1-103.
DR   PDB; 1IIS; X-ray; A/B=107-330.
DR   PDB; 1IIX; X-ray; A/B=107-330.
DR   PDB; 1L6X; X-ray; A=120-326.
DR   PDB; 1OQX; X-ray; A/B=119-330.
DR   PDB; 1T83; X-ray; A/B=107-330.
DR   PDB; 2RCS; X-ray; H=1-103.
DR   HGNC; HGNC:5525; IGHG1.
DR   MIM; 147100; -.
DR   GO; GO:0005624; C:membrane fraction; NAS.
DR   GO; GO:0003823; F:antigen binding; TAS.
DR   GO; GO:0006955; P:immune response; NAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF07654; C1-set; 3.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS00290; IG_MHC; 2.
KW   3D-structure; Direct protein sequencing; Glycoprotein;
KW   Immunoglobulin C region; Immunoglobulin domain.
FT   REGION        1     98       CH1.
FT   REGION       99    110       Hinge.
FT   REGION      111    223       CH2.
FT   REGION      224    330       CH3.
FT   CARBOHYD    180    180       N-linked (GlcNAc. . .).
FT   DISULFID     27     83
FT   DISULFID    103    103       Interchain (with light chain).
FT   DISULFID    109    109       Interchain (with heavy chain).
FT   DISULFID    112    112       Interchain (with heavy chain).
FT   DISULFID    144    204
FT   DISULFID    250    308
FT   VARIANT      97     97       K -> R (in G1M(3) marker).
FT                                /FTId=VAR_003886.
FT   VARIANT     239    239       D -> E (in G1M(non-1) marker).
FT                                /FTId=VAR_003887.
FT   VARIANT     241    241       L -> M (in G1M(non-1) marker).
FT                                /FTId=VAR_003888.
FT   NON_TER       1      1
FT   STRAND       23     24
FT   STRAND       26     33
FT   STRAND       38     38
FT   STRAND       41     41
FT   TURN         42     45
FT   TURN         48     49
FT   STRAND       50     52
FT   STRAND       57     58
FT   TURN         59     61
FT   STRAND       62     71
FT   HELIX        73     75
FT   TURN         76     78
FT   STRAND       82     87
FT   TURN         88     91
FT   STRAND       92     97
FT   TURN        102    103
FT   STRAND      122    126
FT   HELIX       130    134
FT   TURN        136    137
FT   STRAND      141    149
FT   STRAND      157    162
FT   TURN        163    164
FT   STRAND      165    167
FT   STRAND      171    172
FT   STRAND      176    177
FT   TURN        179    180
FT   STRAND      183    190
FT   HELIX       193    197
FT   TURN        198    199
FT   STRAND      202    207
FT   TURN        209    210
FT   STRAND      215    219
FT   STRAND      227    227
FT   STRAND      230    234
FT   HELIX       238    242
FT   STRAND      245    256
FT   STRAND      261    266
FT   TURN        267    268
FT   STRAND      269    270
FT   STRAND      274    276
FT   STRAND      280    281
FT   TURN        283    284
FT   STRAND      287    296
FT   HELIX       297    301
FT   TURN        302    303
FT   STRAND      306    311
FT   TURN        313    314
FT   HELIX       316    318
FT   STRAND      319    324
SQ   SEQUENCE   330 AA;  36106 MW;  3770EE106C2FA33D CRC64;

  Query Match             91.9%;  Score 1233;  DB 1;  Length 330;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 223

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 283

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 330


RESULT 2
Q6GMX6_HUMAN
ID   Q6GMX6_HUMAN PRELIMINARY;      PRT;   465 AA.
AC   Q6GMX6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RA   Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC073766; AAH73766.1; -; mRNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   465 AA;  51083 MW;  B3A9B7D0FDB1386E CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 298

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 358

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 418

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 465


RESULT 3
Q6IN78_HUMAN
ID   Q6IN78_HUMAN PRELIMINARY;      PRT;   466 AA.
AC   Q6IN78;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RG   NIH MGC Project;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072419; AAH72419.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
```

```
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   466 AA;  50854 MW;  53EB0BCEDE81076E CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 466;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 299

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 359

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 419

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        420 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 466


RESULT 4
Q569F4_HUMAN
ID   Q569F4_HUMAN PRELIMINARY;      PRT;   469 AA.
AC   Q569F4;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lymph;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lymph;
RG   NIH MGC Project;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC092518; AAH92518.1; -; mRNA.
SQ   SEQUENCE   469 AA;  51254 MW;  AC13448E3047784F CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 302

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 362

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 422

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        423 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 469


RESULT 5
Q7Z7P5_HUMAN
ID   Q7Z7P5_HUMAN PRELIMINARY;      PRT;   469 AA.
AC   Q7Z7P5;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RG   NIH MGC Project;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC051328; AAH51328.1; -; mRNA.
DR   HSSP; P01857; 1HZH.
DR   SMR; Q7Z7P5; 20-469.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
```

```
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Immunoglobulin domain.
SQ   SEQUENCE   469 AA;  51395 MW;  C8D5BE12BAAF795C CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 302

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 362

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 422

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        423 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 469


RESULT 6
Q7Z5W1_HUMAN
ID   Q7Z5W1_HUMAN PRELIMINARY;      PRT;   470 AA.
AC   Q7Z5W1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RA   Strausberg R.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC053984; AAH53984.1; -; mRNA.
DR   HSSP; P01857; 1HZH.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   470 AA;  51204 MW;  778CF34521483E1A CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 303

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 363

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 423

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        424 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 470


RESULT 7
Q6PJA4_HUMAN
ID   Q6PJA4_HUMAN PRELIMINARY;      PRT;   470 AA.
AC   Q6PJA4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   IGHG1 protein.
GN   Name=IGHG1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Primary B-Cells;
RG   NIH MGC Project;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; BC018747; AAH18747.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   SMR; Q6PJA4; 20-470.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
SQ   SEQUENCE   470 AA;  51716 MW;  7B49556A11FD7D99 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 470;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 303

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 363

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 423

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        424 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 470


RESULT 8
Q6N089_HUMAN
ID   Q6N089_HUMAN PRELIMINARY;      PRT;   472 AA.
AC   Q6N089;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686P15220.
GN   Name=DKFZp686P15220;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Rectum tumor;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640627; CAE45781.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   472 AA;  51724 MW;  26CB340D0046D279 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 305

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 365

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 425

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        426 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 472


RESULT 9
Q5EFE5_HUMAN
ID   Q5EFE5_HUMAN PRELIMINARY;      PRT;   475 AA.
AC   Q5EFE5;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Anti-RhD monoclonal T125 gamma1 heavy chain precursor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Gaucher C., Klein P., Beliard R.;
RT   "Sequence determination of the recombinant human anti-RhD monoclonal
RT   antibody T125.";
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY894992; AAW82028.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   Pfam; PF07686; V-set; 1.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Signal.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20    475       anti-RhD monoclonal T125 gamma1 heavy
FT                                chain.
SQ   SEQUENCE   475 AA;  52362 MW;  1367D400DC7D2859 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 475;
  Best Local Similarity   100.0%;  Pred. No. 5e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 475


RESULT 10
Q6GMW7_HUMAN
ID   Q6GMW7_HUMAN PRELIMINARY;      PRT;   475 AA.
AC   Q6GMW7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RA   Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC073782; AAH73782.1; -; mRNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   475 AA;  51987 MW;  2A1FE55D736860F8 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 475;
  Best Local Similarity   100.0%;  Pred. No. 5e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 475


RESULT 11
Q6GMX1_HUMAN
ID   Q6GMX1_HUMAN PRELIMINARY;      PRT;   476 AA.
AC   Q6GMX1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Spleen;
RA   Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC073773; AAH73773.1; -; mRNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   476 AA;  52286 MW;  622AABA5C62DDE9D CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 476;
```

```
  Best Local Similarity   100.0%;  Pred. No. 5e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 309

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 369

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 429

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        430 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 476
```

```
RESULT 12
Q96PQ8_HUMAN
ID   Q96PQ8_HUMAN PRELIMINARY;      PRT;   679 AA.
AC   Q96PQ8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VII active site mutant immunoconjugate.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21477448; PubMed=11593034; DOI=10.1073/pnas.201420298;
RA   Hu Z., Garen A.;
RT   "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT   cells for immunotherapy in mouse models of prostatic cancer.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Hu Z., Garen A.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF272774; AAK58686.2; -; mRNA.
DR   HSSP; P08709; 1KLI.
DR   SMR; Q96PQ8; 39-180, 191-444, 447-679.
DR   Ensembl; ENSG00000057593; Homo sapiens.
DR   GO; GO:0005576; C:extracellular region; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002383; GLA_blood.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR001254; Peptidase_S1_S6.
DR   InterPro; IPR000294; VitK_dep_GLA.
DR   Pfam; PF07654; C1-set; 2.
DR   Pfam; PF00008; EGF; 1.
DR   Pfam; PF00594; Gla; 1.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PRINTS; PR00001; GLABLOOD.
DR   SMART; SM00179; EGF_CA; 1.
DR   SMART; SM00069; GLA; 1.
DR   SMART; SM00407; IGc1; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS50026; EGF_3; 1.
DR   PROSITE; PS01187; EGF_CA; 1.
DR   PROSITE; PS00011; GLA_1; UNKNOWN_1.
DR   PROSITE; PS50998; GLA_2; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
SQ   SEQUENCE   679 AA;  75552 MW;  0B0023AE70A067A1 CRC64;

  Query Match             91.9%;  Score 1233;  DB 2;  Length 679;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-90;
  Matches  227;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 512

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 572

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 632

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        633 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 679
```

```
RESULT 13
Q6P055_HUMAN
ID   Q6P055_HUMAN PRELIMINARY;      PRT;   473 AA.
AC   Q6P055;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
```

```
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Peripheral Nervous System;
RA   Strausberg R.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC065820; AAH65820.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   473 AA;  51344 MW;  9816D56A77129B57 CRC64;

  Query Match             91.6%;  Score 1229;  DB 2;  Length 473;
  Best Local Similarity   99.6%;  Pred. No. 1e-89;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 306

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 366

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 426

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||| ||||||||||||||||
Db        427 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 473


RESULT 14
Q6MZQ6_HUMAN
ID   Q6MZQ6_HUMAN PRELIMINARY;      PRT;   475 AA.
AC   Q6MZQ6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686G11190.
GN   Name=DKFZp686G11190;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Esophagus tumor;
RG   The German cDNA Consortium;
RA   Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA   Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640947; CAE45972.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   SMR; Q6MZQ6; 20-475.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   475 AA;  52043 MW;  B7EAE255A26F4B8E CRC64;

  Query Match             91.6%;  Score 1229;  DB 2;  Length 475;
  Best Local Similarity   99.6%;  Pred. No. 1e-89;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 308

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 368

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 428

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||| ||||||||||||||||
Db        429 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 475


RESULT 15
Q6N094_HUMAN
ID   Q6N094_HUMAN PRELIMINARY;      PRT;   480 AA.
AC   Q6N094;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein DKFZp686O01196.
GN   Name=DKFZp686O01196;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Esophagus tumor;
RG   The German cDNA Consortium;
RA   Wambutt R., Heubner D., Mewes H.W., Weil B., Amid C., Osanger A.,
RA   Fobo G., Han M., Wiemann S.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX640622; CAE45776.1; -; mRNA.
DR   HSSP; P01861; 1ADQ.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF07654; C1-set; 3.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00407; IGc1; 3.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 4.
DR   PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW   Hypothetical protein.
SQ   SEQUENCE   480 AA;  52612 MW;  225247F3D35AEC18 CRC64;

  Query Match             91.6%;  Score 1229;  DB 2;  Length 480;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-89;
  Matches  226;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         21 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        254 DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 313

Qy         81 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 373

Qy        141 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS 433

Qy        201 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 247
              |||||||||||||||||||||||||| ||||||||||||||||
Db        434 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEGLHNHYTQKSLSLSPGK 480

Search completed: April  4, 2006, 13:15:13
Job time : 190.806 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 3.47251 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-282
Perfect score: 41
Sequence: 1 RPLPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41	100.0	7	2	AAW11128	Aaw11128 Src SH3 d
2	41	100.0	7	2	AAW17010	Aaw17010 SRC SH3 d
3	41	100.0	7	2	AAW25486	Aaw25486 SH3 domai
4	41	100.0	7	2	AAW79781	Aaw79781 Proline-r
5	41	100.0	7	3	AAB17267	Aab17267 Src antag
6	41	100.0	7	3	AAB17226	Aab17226 SH3 antag
7	41	100.0	7	3	AAY69979	Aay69979 Src SH3 r
8	41	100.0	7	4	AAB50762	Aab50762 Human cAM
9	41	100.0	7	5	ABB73219	Abb73219 Src homol
10	41	100.0	7	5	ABB73345	Abb73345 Exemplary
11	41	100.0	7	7	ADJ73499	Adj73499 Exemplary
12	41	100.0	7	7	ADJ73373	Adj73373 SH3 antag
13	41	100.0	7	8	ADJ53007	Adj53007 CH1 delet
14	41	100.0	7	8	ADJ53133	Adj53133 CH1 delet
15	41	100.0	7	8	ADJ51968	Adj51968 CH1 delet
16	41	100.0	7	8	ADJ52094	Adj52094 CH1 delet
17	41	100.0	10	2	AAR93545	Aar93545 Random 10
18	41	100.0	10	2	AAR93544	Aar93544 Random 10
19	41	100.0	10	9	AEB07369	Aeb07369 Signal tr
20	41	100.0	11	3	AAB21130	Aab21130 Src homol
21	41	100.0	11	3	AAB21125	Aab21125 Src homol
22	41	100.0	12	2	AAR93351	Aar93351 FYN prote
23	41	100.0	12	2	AAR93378	Aar93378 Grb-2 pro
24	41	100.0	12	2	AAR93380	Aar93380 Grb-2 pro
25	41	100.0	12	2	AAR93352	Aar93352 FYN prote
26	41	100.0	12	2	AAR93359	Aar93359 LYN prote
27	41	100.0	12	2	AAR93353	Aar93353 FYN prote
28	41	100.0	12	2	AAR93349	Aar93349 FYN prote
29	41	100.0	12	2	AAR93379	Aar93379 Grb-2 pro
30	41	100.0	12	2	AAR93364	Aar93364 LYN prote
31	41	100.0	12	2	AAR93365	Aar93365 LYN prote
32	41	100.0	12	2	AAR93360	Aar93360 LYN prote
33	41	100.0	12	2	AAR93362	Aar93362 LYN/PI3K
34	41	100.0	12	2	AAR93363	Aar93363 LYN prote
35	41	100.0	12	2	AAR93343	Aar93343 SRC prote
36	41	100.0	12	2	AAR93344	Aar93344 SRC prote
37	41	100.0	12	2	AAR93346	Aar93346 SRC/LYN p
38	41	100.0	12	2	AAR93377	Aar93377 Grb-2 pro
39	41	100.0	12	2	AAR93345	Aar93345 SRC prote
40	41	100.0	12	2	AAR93348	Aar93348 SRC prote
41	41	100.0	12	2	AAR93356	Aar93356 FYN/LYN p
42	41	100.0	12	3	AAB17254	Aab17254 SH3 antag
43	41	100.0	12	3	AAB17251	Aab17251 SH3 antag
44	41	100.0	12	3	AAB17255	Aab17255 SH3 antag
45	41	100.0	12	3	AAB17253	Aab17253 SH3 antag

ALIGNMENTS

RESULT 1
AAW11128
ID AAW11128 standard; peptide; 7 AA.
XX
AC AAW11128;
XX
DT 27-JUN-1997 (first entry)
XX
DE Src SH3 domain-binding peptide preferred core sequence.
XX
KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
OS Synthetic.
XX
PN WO9603649-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US009382.
XX
PR 22-JUL-1994; 94US-00278865.
PR 07-JUN-1995; 95US-00483555.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
XX
DR WPI; 1996-117151/12.
XX
PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.
XX
PS Disclosure; Page 62; 116pp; English.
XX
CC AAW11128 represents a preferred core sequence of a set of peptides that
CC bind to the Src SH3 domain. The SH3 binding peptides are useful in
CC modulating signal transduction pathways at the cellular level (especially
CC protein tyrosine kinase-mediated), modulating oncogenic protein activity,
CC or providing compounds for the development of drugs with the ability to
CC modulate broad classes, as well as specific classes, of proteins involved
CC in signal transduction and also for regulating the processing,
CC trafficking or translation of RNA. Conjugates of the peptides with
CC detectable labels or imaging agents are useful for imaging cells, tissues
CC and organs in which Src or Src-related proteins are expressed

```
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db          1 RPLPPLP 7


RESULT 2
AAW17010
ID   AAW17010 standard; peptide; 7 AA.
XX
AC   AAW17010;
XX
DT   27-JUN-1997  (first entry)
XX
DE   SRC SH3 domain-binding consensus peptide.
XX
KW   Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW   protein tyrosine kinase; signal transduction; RNA processing;
KW   trafficking; translation.
XX
OS   Synthetic.
XX
PN   WO9603649-A1.
XX
PD   08-FEB-1996.
XX
PF   24-JUL-1995;   95WO-US009382.
XX
PR   22-JUL-1994;   94US-00278865.
PR   07-JUN-1995;   95US-00483555.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ;
XX
DR   WPI; 1996-117151/12.
XX
PT   Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT   proteins - useful for e.g. modulating signal transduction pathways at the
PT   cellular level, esp. protein tyrosine kinase-mediated.
XX
PS   Example 14; Page 58; 116pp; English.
XX
CC   AAW17010 is the consensus sequence of a set of SRC SH3-binding peptides
CC   derived from a biased peptide library, exhibiting selective SH3 binding.
CC   SH3 binding peptides are useful in modulating signal transduction
CC   pathways at the cellular level (especially protein tyrosine kinase-
CC   mediated), modulating oncogenic protein activity, or providing compounds
CC   for the development of drugs with the ability to modulate broad classes,
CC   as well as specific classes, of proteins involved in signal transduction
CC   and also for regulating the processing, trafficking or translation of
CC   RNA. Conjugates of the peptides with detectable labels or imaging agents
CC   are useful for imaging cells, tissues and organs in which Src or Src-
CC   related proteins are expressed
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db          1 RPLPPLP 7


RESULT 3
AAW25486
ID   AAW25486 standard; peptide; 7 AA.
XX
AC   AAW25486;
XX
DT   27-MAR-1998  (first entry)
XX
DE   SH3 domain binding peptide consensus motif.
XX
KW   Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW   tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW   PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS   Synthetic.
OS   Unidentified.
XX
PN   WO9730074-A1.
XX
PD   21-AUG-1997.
XX
PF   14-FEB-1997;   97WO-US002298.
XX
PR   16-FEB-1996;   96US-00602999.
XX
PA   (CYTO-) CYTOGEN CORP.
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ,  Fowlkes DM;
PI   Rider JE;
XX
DR   WPI; 1997-424972/39.
XX
PT   Src homology region 3 binding peptide - used to activate Src tyrosine
PT   kinase(s) and to stimulate immune response by increasing production of
PT   certain lymphokine(s), e.g. interleukin-1.
XX
PS   Disclosure; Page 8; 131pp; English.
XX
CC   The present sequence represents a Src homology region 3 (SH3) binding
CC   peptide consensus motif. SH3 binding peptides are selected from: (a)
CC   peptides which bind the SH3 domain of Cortactin; (b) peptides which bind
CC   the middle SH3 domain of Nck; (c) peptides which bind the SH3 domain of
CC   Abl; (d) peptides which bind the SH3 domain of Src; (e) peptides which
CC   bind the SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain
CC   of p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk;
CC   (h) peptides which bind the SH3 domain of Yes; and (i) peptides which
CC   bind the amino-terminal SH3 domain of Grb2. The purified binding peptides
CC   can be used in the method to identify inhibitors of their binding to
CC   their respective SH3 domains, which could be used to modulate the
CC   pharmacological activity of proteins or polypeptide containing the SH3
CC   domain. The peptides can also be used to activate Src or Src-related
CC   protein tyrosine kinases, to stimulate the immune response by increasing
CC   the production of certain lymphokines, e.g. tumour necrosis factor-alpha
CC   and interleukin-1, or to deliver a conjugated molecule to certain
CC   cellular compartments containing Src or Src related proteins
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db          1 RPLPPLP 7


RESULT 4
AAW79781
ID   AAW79781 standard; peptide; 7 AA.
XX
AC   AAW79781;
XX
```

DT 18-JAN-1999 (first entry)
XX
DE Proline-rich peptide which binds with Src SH3 domain.
XX
KW electrochemical; potentiometric; specific binding pair; assay;
KW competition; analysis; purification; proline-rich; Src SH3.
XX
OS Unidentified.
XX
PN WO9835232-A2.
XX
PD 13-AUG-1998.
XX
PF 06-FEB-1998; 98WO-US002440.
XX
PR 06-FEB-1997; 97US-0036919P.
PR 16-SEP-1997; 97US-0059049P.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
PA (NOVA-) NOVALON PHARM CORP.
XX
PI Fowlkes DM, Thorp HH;
XX
DR WPI; 1998-467163/40.
XX
PT Apparatus for electrochemically detecting binding - for use in
PT biochemical analyses, drug development and protein purification assays.
XX
PS Example 6; Page 55; 104pp; English.
XX
CC The invention relates to a method and apparatus for performing an
CC electrochemical assay for detecting specific binding between members of a
CC biological binding pair. The apparatus detects specific binding between a
CC first member immobilised on an electrode and a second member which is
CC biologically labelled, in the presence of an electrochemical mediator.
CC The method may be used for performing binding and competitive binding
CC assays. It may be used in performing high throughput screening assays for
CC detecting inhibition of specific binding between the members of the
CC binding pair for use in drug development, biochemical analyses and
CC protein purification assays. The present sequence is an example of a
CC peptide which is used in labelled form as a second binding member in the
CC above assay. The peptide acts as a surrogate ligand for the the first
CC member. Specifically, the peptide is a proline-rich peptide which binds
CC with Src SH3 domain
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 5
AAB17267
ID AAB17267 standard; peptide; 7 AA.
XX
AC AAB17267;
XX
DT 31-OCT-2000 (first entry)
XX
DE Src antagonist peptide sequence SEQ ID NO:323.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 308; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 6
AAB17226
ID AAB17226 standard; peptide; 7 AA.
XX
AC AAB17226;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:282.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX

PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 295; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 7
AAY69979
ID AAY69979 standard; peptide; 7 AA.
XX
AC AAY69979;
XX
DT 14-APR-2000 (first entry)
XX
DE Src SH3 region binding peptide #1.
XX
KW Electrochemically labelled peptide; probe; electrochemical assay;
KW binding detection; biological binding pair; electrochemical analysis;
KW drug detection; drug development; biochemical analysis; Src;
KW protein purification assay; SH3 region.
XX
OS Synthetic.
XX
PN WO9964847-A1.
XX
PD 16-DEC-1999.
XX
PF 28-MAY-1999; 99WO-US011848.
XX
PR 08-JUN-1998; 98US-00093444.
XX
PA (XANT-) XANTHON INC.
XX
PI Welch TW;
XX
DR WPI; 2000-136855/12.
XX
PT Apparatus for electrochemical analyses for drug detection, etc.
XX
PS Example 5; Page 46; 104pp; English.
XX
CC This sequence represents an Src SH3 region binding peptide. The invention
CC relates to an apparatus for performing an electrochemical assay for
CC detecting binding between members of a biological binding pair. The
CC apparatus has: a first electrode (comprising a conducting or
CC semiconducting surface); a second, reference electrode (comprising a
CC conducting metal in contact with an aqueous electrolyte solution); and a
CC third, auxiliary electrode; where each electrode is connected to a
CC potentiostat and is in contact with an electrolyte solution containing
CC both members of a biological binding pair. The second member of the
CC binding pair is electrochemically labeled with a chemical species capable
CC of participating in a reduction/oxidation reaction at the surface of the
CC first electrode under conditions where an electrical potential is applied
CC to the electrodes. A current is produced in the apparatus when an
CC electrical potential is applied to the electrodes, and the current is
CC reduced upon binding of the second member of the biological binding pair
CC to the first member of the pair. The method can be used for
CC electrochemical analyses for drug detection, drug development,
CC biochemical analysis and protein purification assays. The method provides
CC a means for rapid, high throughput screening of biologically active
CC compounds
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 8
AAB50762
ID AAB50762 standard; peptide; 7 AA.
XX
AC AAB50762;
XX
DT 20-MAR-2001 (first entry)
XX
DE Human cAMP-specific phosphodiesterase PDE4D5 modulator SEQ ID NO: 26.
XX
KW PDE4D5; cAMP-specific phosphodiesterase; RACK1; modulator;
KW receptor for activated C-kinase.
XX
OS Unidentified.
XX
PN WO200071080-A2.
XX
PD 30-NOV-2000.
XX
PF 20-MAY-2000; 2000WO-US013961.
XX
PR 20-MAY-1999; 99US-0135035P.
XX
PA (UTAH) UNIV UTAH RES FOUND.
XX
PI Bolger GB, Houslay MD, Steele MR, Yarwood SJ;
XX
DR WPI; 2001-061280/07.
XX

PT Screening drugs that modulate activity of cAMP-specific phosphodiesterase
PT for treating various conditions by detecting modulation of interaction
PT between phosphodiesterase and activated C-kinase receptor by the drug.
XX
PS Example 2; Page 43; 77pp; English.
XX
CC The present invention provides methods and peptides for use in
CC identifying modulators of the cAMP-specific phosphodiesterase isoform
CC PDE4D5. These act by modulating the interaction of PDE4D5 with the
CC receptor for activated C-kinase (RACK1). The modulators are useful in the
CC treatment of various conditions
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 9
ABB73219
ID ABB73219 standard; peptide; 7 AA.
XX
AC ABB73219;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:282.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 10
ABB73345
ID ABB73345 standard; peptide; 7 AA.
XX
AC ABB73345;
XX
DT 05-APR-2002 (first entry)
XX
DE Exemplary pharmacologically active peptide SEQ ID NO:323.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX

PS Claim 39; Page 61; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 11
ADJ73499
ID ADJ73499 standard; peptide; 7 AA.
XX
AC ADJ73499;
XX
DT 06-MAY-2004 (first entry)
XX
DE Exemplary mimetic peptide sequence SeqID 955.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 955; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an exemplary mimetic peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 12
ADJ73373
ID ADJ73373 standard; peptide; 7 AA.
XX
AC ADJ73373;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 827.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 827; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human

CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 13
ADJ53007
ID ADJ53007 standard; peptide; 7 AA.
XX
AC ADJ53007;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID827.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 827; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 14
ADJ53133
ID ADJ53133 standard; peptide; 7 AA.
XX
AC ADJ53133;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID955.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 955; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous

CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

RESULT 15
ADJ51968
ID ADJ51968 standard; peptide; 7 AA.
XX
AC ADJ51968;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID827.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 827; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 1 RPLPPLP 7

Search completed: April 4, 2006, 13:07:45
Job time : 4.47251 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 1.14529 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-282
Perfect score: 41
Sequence: 1 RPLPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41	100.0	339	2	JC7509	glycoprotein VI-1
2	41	100.0	459	2	S03116	gene 33 protein, h
3	41	100.0	477	2	T46917	hypothetical prote
4	41	100.0	535	2	A46101	protein-tyrosine-p
5	41	100.0	548	2	B46101	protein-tyrosine-p
6	41	100.0	613	2	A56031	potassium channel
7	41	100.0	663	1	TVMVRR	protein-tyrosine k
8	41	100.0	894	2	F84870	hypothetical prote
9	39	95.1	173	2	T19341	hypothetical prote
10	39	95.1	409	2	S60975	hypothetical prote
11	39	95.1	514	2	C49507	potassium channel
12	39	95.1	602	2	A49507	potassium channel
13	38	92.7	222	2	C75539	conserved hypothet
14	38	92.7	527	2	I84483	protein-tyrosine k
15	38	92.7	527	2	I49133	protein-tyrosine k
16	38	92.7	668	2	S56909	polymyxin B resist
17	36	87.8	104	1	R5PM25	ribosomal protein
18	36	87.8	123	2	AH2707	conserved hypothet
19	36	87.8	162	2	T07173	hypothetical prote
20	36	87.8	169	2	H72470	hypothetical prote
21	36	87.8	196	2	T00702	hypothetical prote
22	36	87.8	198	2	E89008	protein W08A12.3 [
23	36	87.8	238	2	T32889	hypothetical prote
24	36	87.8	256	2	A35340	H+-transporting tw
25	36	87.8	265	2	H75560	oxidoreductase, sh
26	36	87.8	277	2	I38857	microtubule-associ
27	36	87.8	281	2	I38707	Fas ligand - human
28	36	87.8	291	2	G84494	hypothetical prote
29	36	87.8	341	2	E69463	type I restriction
30	36	87.8	348	2	D88088	protein B0454.1 [i
31	36	87.8	378	2	T28112	hypothetical prote
32	36	87.8	417	2	G64417	hypothetical prote
33	36	87.8	426	2	F95058	hypothetical prote
34	36	87.8	428	2	S76184	hypothetical prote
35	36	87.8	431	2	S20065	nuclear factor I-X
36	36	87.8	522	2	H97927	type 1 site-specif
37	36	87.8	524	2	A75588	probable protein k
38	36	87.8	598	2	T02265	hypothetical prote
39	36	87.8	604	2	H81110	sulfite reductase
40	36	87.8	606	1	UZAD12	terminal protein p
41	36	87.8	609	2	T28736	hypothetical prote
42	36	87.8	627	2	T26064	hypothetical prote
43	36	87.8	645	2	T16078	hypothetical prote
44	36	87.8	653	1	UZADP2	terminal protein p
45	36	87.8	653	1	UZADP5	terminal protein p

ALIGNMENTS

RESULT 1
JC7509
glycoprotein VI-1 - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C;Accession: JC7509; PC7101
R;Ezumi, Y.; Uchiyama, T.; Takayama, H.
Biochem. Biophys. Res. Commun. 277, 27-36, 2000
A;Title: Molecular cloning, genomic structure, chromosomal localization, and alternati
A;Reference number: JC7509; MUID:20483673; PMID:11027634
A;Contents: Platelet
A;Accession: JC7509
A;Molecule type: mRNA
A;Residues: 1-339 <EZU>
A;Cross-references: UNIPROT:Q9UIF2; UNIPARC:UPI000006F4A8; DDBJ:AB043819
A;Accession: PC7101
A;Molecule type: protein
A;Residues: 28-41;62-79;114-142 <EZ2>
A;Cross-references: UNIPARC:UPI000017A509; UNIPARC:UPI000017A50A; UNIPARC:UPI000017A5C
C;Comment: This protein, which belongs to the immunoglobulin superfamily, is the major
or gamma chain as a signal transducing subunit, and plays some roles in cancer cells.
C;Genetics:
A;Gene: gpVI-1
A;Map position: 19q13.4
A;Introns: 62/1; 95/1; 353/1; 638/1; 692/1; 752/1; 803/1
C;Keywords: glycoprotein; immunoglobulin; platelet

Query Match 100.0%; Score 41; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPPLP 7
              |||||||
Db        307 RPLPPLP 313
```

RESULT 2
S03116
gene 33 protein, hepatic - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S03116; S03402; B30568; S08283
R;Chrapkiewicz, N.B.; Davis, C.M.; Chu, D.T.W.; Caldwell, C.M.; Granner, D.K.
Nucleic Acids Res. 17, 6651-6667, 1989
A;Title: Rat gene 33: analysis of its structure, messenger RNA and basal promoter acti
A;Reference number: S03116; MUID:89385990; PMID:2780291
A;Accession: S03116
A;Molecule type: DNA
A;Residues: 1-459 <CHR>
A;Cross-references: UNIPROT:P05432; UNIPARC:UPI000012F0FD; EMBL:X07266; NID:g57568; PI
R;Lee, K.L.; Makkinje, A.; Ch'ang, L.Y.; Kenney, F.T.
Arch. Biochem. Biophys. 269, 106-113, 1989

A;Title: Molecular cloning and analysis of full-length cDNAs cognate to a rat gene under
A;Reference number: S03402; MUID:89133523; PMID:2916834
A;Accession: S03402
A;Molecule type: mRNA
A;Residues: 1-459 <LEE>
A;Cross-references: UNIPARC:UPI000012F0FD
A;Note: the authors translated the codon GGA for residue 18 as Lys, TAC for residues 192
A;Accession: B30568
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-17,'K',19-66,143-191,'T',193-301,'L',303-310,'L',312-395,'L',397-409,'L',4
A;Cross-references: UNIPARC:UPI000017C912
R;Lee, K.L.; Makkinje, A.; Ch'ang, L.Y.; Kenney, F.T.
Arch. Biochem. Biophys. 276, 554, 1990
A;Reference number: S08283
A;Contents: annotation
A;Note: this is a revision of the nucleotide translation of residues 18, 192, 302, 311,
C;Genetics:
A;Gene: 33
C;Keywords: alternative splicing; liver
F;1-459/Product: gene 33 protein, long form #status predicted <MAT1>
F;1-66,143-459/Product: gene 33 protein, short form #status predicted <MAT2>

Query Match 100.0%; Score 41; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 150 RPLPPLP 156

RESULT 3
T46917
hypothetical protein DKFZp762K137.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: T46917
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24136
A;Accession: T46917
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-477 <AAA>
A;Cross-references: UNIPROT:Q9NSQ8; UNIPARC:UPI000006CEC3; EMBL:AL157482
A;Experimental source: adult melanoma (MeWo cell line); clone DKFZp762K137
C;Genetics:
A;Note: DKFZp762K137.1

Query Match 100.0%; Score 41; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 228 RPLPPLP 234

RESULT 4
A46101
protein-tyrosine-phosphatase (EC 3.1.3.48) nonreceptor type PTP61F, short splice form -
C;Species: Drosophila melanogaster
C;Date: 08-May-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: A46101
R;McLaughlin, S.; Dixon, J.E.
J. Biol. Chem. 268, 6839-6842, 1993
A;Title: Alternative splicing gives rise to a nuclear protein tyrosine phosphatase in Dr
A;Reference number: A46101; MUID:93216607; PMID:8463208
A;Accession: A46101
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-535 <MCL>
A;Cross-references: UNIPROT:Q9W0G1; UNIPARC:UPI000016BD13; GB:L11251; NID:g290265; PIDN
A;Note: authors translated the codon TTC for residue 382 as Ile, and CGA for residue 52
C;Genetics:
A;Gene: FlyBase:Ptp61F
A;Cross-references: FlyBase:FBgn0003138
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type PTP61F; protein-tyrosine-
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyros
F;60-285/Domain: protein-tyrosine-phosphatase homology <PTP>
F;237/Active site: Cys (phosphocysteine intermediate) #status predicted
F;243/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 41; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 391 RPLPPLP 397

RESULT 5
B46101
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type PTP61F, long splice form -
C;Species: Drosophila melanogaster
C;Date: 08-May-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C;Accession: B46101
R;McLaughlin, S.; Dixon, J.E.
J. Biol. Chem. 268, 6839-6842, 1993
A;Title: Alternative splicing gives rise to a nuclear protein tyrosine phosphatase in D
A;Reference number: A46101; MUID:93216607; PMID:8463208
A;Accession: B46101
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-548 <MCL>
A;Cross-references: UNIPROT:Q9W0G1; UNIPARC:UPI000016BD12; GB:L11251
A;Note: authors translated the codon TTC for residue 382 as Ile
C;Genetics:
A;Gene: FlyBase:Ptp61F
A;Cross-references: FlyBase:FBgn0003138
C;Superfamily: protein-tyrosine-phosphatase, nonreceptor type PTP61F; protein-tyrosine-
C;Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyros
F;60-285/Domain: protein-tyrosine-phosphatase homology <PTP>
F;237/Active site: Cys (phosphocysteine intermediate) #status predicted
F;243/Binding site: substrate phosphate (Arg) #status predicted

Query Match 100.0%; Score 41; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||||
Db 391 RPLPPLP 397

RESULT 6
A56031
potassium channel KCNA5 - human
N;Alternate names: potassium channel HK2; potassium channel PCN1; shaker-related potas
C;Species: Homo sapiens (man)
C;Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C;Accession: A56031; A38556; B39922; A38074
R;Philipson, L.H.; LaMendola, J.; Bell, G.I.; Steiner, D.F.
submitted to GenBank, September 1990
A;Reference number: A56031
A;Accession: A56031
A;Molecule type: mRNA
A;Residues: 1-613 <PHI>
A;Cross-references: UNIPROT:P22460; UNIPARC:UPI000016AE76; GB:M55513; NID:g189653; PID
R;Philipson, L.H.; Hice, R.E.; Schaefer, K.; LaMendola, J.; Bell, G.I.; Nelson, D.J.;
Proc. Natl. Acad. Sci. U.S.A. 88, 53-57, 1991
A;Title: Sequence and functional expression in Xenopus oocytes of a human insulinoma a
A;Reference number: A38556; MUID:91095456; PMID:1986382
A;Accession: A38556

```
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-56,'G',58-137,'L',139-213,'R',215-227,'P',229-545,'QG',546-613 <PH2>
A;Cross-references: UNIPARC:UPI00001779F3; GB:M55513
R;Tamkun, M.M.; Knoth, K.M.; Walbridge, J.A.; Kroemer, H.; Roden, D.M.; Glover, D.M.
FASEB J. 5, 331-337, 1991
A;Title: Molecular cloning and characterization of two voltage-gated K(+) channel cDNAs
A;Reference number: A39922; MUID:91160866; PMID:2001794
A;Accession: B39922
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-54,56-137,'L',139-186,'G',189-213,'R',215-227,'P',229-297,'PTQRGH',309-558
A;Cross-references: UNIPARC:UPI00001779F4; GB:M60451
R;Curran, M.E.; Landes, G.M.; Keating, M.T.
Genomics 12, 729-737, 1992
A;Title: Molecular cloning, characterization, and genomic localization of a human potass
A;Reference number: A38074; MUID:92241872; PMID:1349297
A;Accession: A38074
A;Molecule type: DNA
A;Residues: 1-137,'L',139-153,'R',155-213,'R',215-227,'P',229-281,'V',283-578,'QLPPREV'
A;Cross-references: UNIPARC:UPI000016AEEA; GB:M83254; NID:g190202; PIDN:AAA60146.1; PID:
A;Experimental source: heart
A;Note: sequence extracted from NCBI backbone (NCBIN:98573, NCBIP:98577)
C;Genetics:
A;Gene: GDB:KCNA5
A;Cross-references: GDB:127904; OMIM:176267
A;Map position: 12p13.33-12p13.31
C;Superfamily: potassium channel protein drk1
C;Keywords: glycoprotein; phosphoprotein; potassium channel; transmembrane protein; volt
F;125,190/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;557/Binding site: phosphate (Ser) (covalent) #status predicted

  Query Match             100.0%;  Score 41;  DB 2;  Length 613;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db         65 RPLPPLP 71


RESULT 7
TVMVRR
protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rashee
C;Species: feline sarcoma virus
A;Note: host Felis sp. (cat)
C;Date: 27-Nov-1985 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: A00653; A03937
R;Naharro, G.; Robbins, K.C.; Reddy, E.P.
Science 223, 63-66, 1984
A;Title: Gene product of v-fgr onc: hybrid protein containing a portion of actin and a t
A;Reference number: A00653; MUID:84097512; PMID:6318314
A;Accession: A00653
A;Molecule type: DNA
A;Residues: 1-663 <NAH>
A;Cross-references: UNIPROT:P00544; UNIPARC:UPI000017101E; GB:X00255; GB:K01487; NID:g61
A;Note: the authors translated the codon GAT for residue 14 as Glu
C;Comment: This protein is synthesized as a gag-fgr polyprotein.
C;Genetics:
A;Gene: fgr
C;Superfamily: feline sarcoma virus protein-tyrosine kinase fgr; protein kinase homology
C;Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; poly
F;1-118/Region: gag polyprotein similarity
F;141-268/Region: actin similarity
F;285-382/Domain: SH2 homology <SH2>
F;402-660/Domain: protein kinase homology <KIN>
F;410-418/Region: protein kinase ATP-binding motif
F;432/Active site: Lys #status predicted

  Query Match             100.0%;  Score 41;  DB 1;  Length 663;
  Best Local Similarity   100.0%;  Pred. No. 63;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        131 RPLPPLP 137


RESULT 8
F84870
hypothetical protein At2g43800 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84870
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84870
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-894 <STO>
A;Cross-references: UNIPROT:O22824; UNIPARC:UPI00000A2AE1; GB:AE002093; NID:g2281090;
C;Genetics:
A;Gene: At2g43800
A;Map position: 2

  Query Match             100.0%;  Score 41;  DB 2;  Length 894;
  Best Local Similarity   100.0%;  Pred. No. 86;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        289 RPLPPLP 295


RESULT 9
T19341
hypothetical protein C16D6.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19341
R;Gardner, A.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19111
A;Accession: T19341
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-173 <WIL>
A;Cross-references: UNIPROT:O62063; UNIPARC:UPI000007A889; EMBL:Z81472; PIDN:CAB03889.
A;Experimental source: clone C16D6
C;Genetics:
A;Gene: CESP:C16D6.1
A;Map position: X
A;Introns: 42/1

  Query Match             95.1%;  Score 39;  DB 2;  Length 173;
  Best Local Similarity   85.7%;  Pred. No. 31;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              ||:||||
Db        121 RPMPPLP 127


RESULT 10
S60975
hypothetical protein YNL152w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein N1765
C;Species: Saccharomyces cerevisiae
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C;Accession: S60975; S63104; S63822
R;Nasr, F.; Becam, A.M.; Herbert, C.J.
submitted to the EMBL Data Library, October 1995
```

A;Description: The sequence of 36.8 kb from the left arm of chromosome XIV reveals 24 cc
tonic dystrophy kinase.
A;Reference number: S60958
A;Accession: S60975
A;Molecule type: DNA
A;Residues: 1-409 <NAS>
A;Cross-references: UNIPROT:P53901; UNIPARC:UPI000013BB88; EMBL:X92517; NID:g1050783; PI
R;Nasr, F.; Becam, A.M.; Herbert, C.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S62967
A;Accession: S63104
A;Molecule type: DNA
A;Residues: 1-409 <NAW>
A;Cross-references: UNIPARC:UPI000013BB88; EMBL:Z71428; NID:g1302109; PID:e239813; PID:g
A;Experimental source: strain S288C
R;Nasr, F.; Becam, A.M.; Herbert, C.J.
Yeast 12, 169-175, 1996
A;Title: The sequence of 36.8 kb from the left arm of chromosome XIV reveals 24 complete
dystrophy kinase.
A;Reference number: S63805; MUID:96287653; PMID:8686380
A;Accession: S63822
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-409 <NAF>
A;Cross-references: UNIPARC:UPI000013BB88; EMBL:X92517; NID:g1050783; PIDN:CAA63287.1; P
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Cross-references: SGD:S0005096
A;Map position: 14L
A;Note: YNL152w

  Query Match 95.1%; Score 39; DB 2; Length 409;
  Best Local Similarity 85.7%; Pred. No. 76;
  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
||:||||
Db 159 RPIPPLP 165

RESULT 11
C49507
potassium channel Kv1.5, form 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 17-Nov-2000
C;Accession: C49507
R;Attali, B.; Lesage, F.; Ziliani, P.; Guillemare, E.; Honore, E.; Waldmann, R.; Hugnot,
J. Biol. Chem. 268, 24283-24289, 1993
A;Title: Multiple mRNA isoforms encoding the mouse cardiac Kv1-5 delayed rectifier K(+)
A;Reference number: A49507; MUID:94043264; PMID:8226976
A;Accession: C49507
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-514 <ATT>
A;Cross-references: UNIPARC:UPI00001779F5; GB:L22218
C;Superfamily: potassium channel protein drk1
C;Keywords: alternative splicing

  Query Match 95.1%; Score 39; DB 2; Length 514;
  Best Local Similarity 85.7%; Pred. No. 96;
  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||:|
Db 65 RPLPPMP 71

RESULT 12
A49507
potassium channel Kv1.5 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: A49507; B49507
R;Attali, B.; Lesage, F.; Ziliani, P.; Guillemare, E.; Honore, E.; Waldmann, R.; Hugnot,
J. Biol. Chem. 268, 24283-24289, 1993
A;Title: Multiple mRNA isoforms encoding the mouse cardiac Kv1-5 delayed rectifier K(+)
A;Reference number: A49507; MUID:94043264; PMID:8226976
A;Accession: A49507
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-602 <ATT>
A;Cross-references: UNIPROT:Q61762; UNIPARC:UPI0000028EB9; GB:L22218; NID:g435603; PIDN
A;Accession: B49507
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 201-602 <AT2>
A;Cross-references: UNIPARC:UPI000002A65A; GB:L22218
C;Superfamily: potassium channel protein drk1
C;Keywords: alternative splicing

  Query Match 95.1%; Score 39; DB 2; Length 602;
  Best Local Similarity 85.7%; Pred. No. 1.1e+02;
  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
|||||:|
Db 65 RPLPPMP 71

RESULT 13
C75539
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: C75539
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: C75539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <WHI>
A;Cross-references: UNIPROT:Q9RXN1; UNIPARC:UPI00000C171C; GB:AE001889; GB:AE000513; NI
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0279
A;Map position: 1

  Query Match 92.7%; Score 38; DB 2; Length 222;
  Best Local Similarity 85.7%; Pred. No. 56;
  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
||:||||
Db 201 RPVPPLP 207

RESULT 14
I84483
protein-tyrosine kinase (EC 2.7.1.112) TXK - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Oct-2004
C;Accession: I84483; I38372; I38373; I38374; I38375
R;Haire, R.N.; Ohta, Y.; Lewis, J.E.; Fu, S.M.; Kroisel, P.; Litman, G.W.
Hum. Mol. Genet. 3, 897-901, 1994
A;Title: TXK, a novel human tyrosine kinase expressed in T cells shares sequence identi
A;Reference number: I38372; MUID:95038742; PMID:7951233
A;Accession: I84483
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-527 <HAR1>
A;Cross-references: UNIPROT:P42681; UNIPARC:UPI0000137828; GB:L27071; NID:g951045; PIDN
A;Accession: I38372

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 7-24 <HAR2>
A;Cross-references: UNIPARC:UPI000000052F; EMBL:U07791; NID:g508216; PIDN:AAA19597.1; PI
A;Accession: I38373
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 128-148 <HAR3>
A;Cross-references: UNIPARC:UPI0000000530; EMBL:U07792; NID:g508217; PIDN:AAA19598.1; PI
A;Accession: I38374
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 195-236 <HAR4>
A;Cross-references: UNIPARC:UPI000011DE8E; EMBL:U07793; NID:g508218; PIDN:AAA19599.1; PI
A;Accession: I38375
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 238-275,'X',277-318 <HAR5>
A;Cross-references: UNIPARC:UPI000011DDEE; EMBL:U07794; NID:g508219; PIDN:AAA19600.1; PI
C;Genetics:
A;Gene: GDB:TXK
A;Cross-references: GDB:377329; OMIM:600058
A;Map position: 4p12-4p12
A;Introns: 262/1
C;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F;89-137/Domain: SH3 homology <SH3>
F;150-246/Domain: SH2 homology <SH2>
F;269-527/Domain: protein kinase homology <KIN>
F;299/Active site: Lys #status predicted

Query Match 92.7%; Score 38; DB 2; Length 527;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
:||||||
Db 70 KPLPPLP 76

RESULT 15
I49133
protein-tyrosine kinase (EC 2.7.1.112) txk - mouse
N;Alternate names: resting lymphocyte protein-tyrosine kinase (rlk)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Oct-2004
C;Accession: I49133; A55631
R;Haire, R.N.; Litman, G.W.
Mamm. Genome 6, 476-480, 1995
A;Title: The murine form of TXK, a novel TEC kinase expressed in thymus maps to chromosc
A;Reference number: I49133; MUID:96059536; PMID:7579892
A;Accession: I49133
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-527 <HAI>
A;Cross-references: UNIPROT:P42682; UNIPARC:UPI0000028034; EMBL:U16145; NID:g562124; PID
R;Hu, Q.; Davidson, D.; Schwartzberg, P.L.; Macchiarini, F.; Lenardo, M.J.; Bluestone, J
J. Biol. Chem. 270, 1928-1934, 1995
A;Title: Identification of rlk, a novel protein tyrosine kinase with predominant express
A;Reference number: A55631; MUID:95130578; PMID:7829530
A;Accession: A55631
A;Molecule type: mRNA
A;Residues: 1-527 <HUA>
A;Cross-references: UNIPARC:UPI0000028034; GB:L35268; NID:g623442; PIDN:AAA67039.1; PID:
A;Note: in Genbank entry MUSRLK, release 116.0, the source is designated as Mus cookii
C;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F;89-137/Domain: SH3 homology <SH3>
F;150-246/Domain: SH2 homology <SH2>
F;269-527/Domain: protein kinase homology <KIN>
F;277-285/Region: protein kinase ATP-binding motif

Query Match 92.7%; Score 38; DB 2; Length 527;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPPLP 7
:||||||
Db 71 KPLPPLP 77

Search completed: April 4, 2006, 13:17:28
Job time : 2.14529 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 5.35079 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-282
Perfect score: 41
Sequence: 1 RPLPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41	100.0	40	2	Q62KF6_BURMA	Q62kf6 burkholderi
2	41	100.0	101	2	Q8QS70_9BETA	Q8qs70 pongine her
3	41	100.0	128	2	Q8STT3_ENCCU	Q8stt3 encephalito
4	41	100.0	168	2	Q67TQ3_ORYSA	Q67tq3 oryza sativ
5	41	100.0	199	2	Q7XR59_ORYSA	Q7xr59 oryza sativ
6	41	100.0	271	2	Q80WJ0_MOUSE	Q80wj0 mus musculu
7	41	100.0	279	2	Q6N908_RHOPA	Q6n908 rhodopseudo
8	41	100.0	282	1	TNFL6_PIG	Q9bea8 sus scrofa
9	41	100.0	321	2	Q9HCN7_HUMAN	Q9hcn7 homo sapien
10	41	100.0	339	2	Q9UIF2_HUMAN	Q9uif2 homo sapien
11	41	100.0	378	2	Q5YME9_NOCFA	Q5yme9 nocardia fa
12	41	100.0	401	2	Q6PA50_XENLA	Q6pa50 xenopus lae
13	41	100.0	404	2	Q6DF84_XENLA	Q6df84 xenopus lae
14	41	100.0	423	2	Q872X2_NEUCR	Q872x2 neurospora
15	41	100.0	431	2	Q8IRH4_DROME	Q8irh4 drosophila
16	41	100.0	443	2	Q67VV7_ORYSA	Q67vv7 oryza sativ
17	41	100.0	459	1	MIG6_RAT	P05432 rattus norv
18	41	100.0	461	1	MIG6_MOUSE	Q99jz7 mus musculu
19	41	100.0	462	1	MIG6_HUMAN	Q9ujm3 homo sapien
20	41	100.0	477	2	Q9NSQ8_HUMAN	Q9nsq8 homo sapien
21	41	100.0	488	2	Q51A29_ENTHI	Q51a29 entamoeba h
22	41	100.0	495	2	Q9H5I7_HUMAN	Q9h5i7 homo sapien
23	41	100.0	499	2	Q51A34_ENTHI	Q51a34 entamoeba h
24	41	100.0	526	2	Q501W1_RAT	Q501w1 rattus norv
25	41	100.0	545	1	FGR_FSVGR	P00544 feline sarc
26	41	100.0	548	1	PTP61_DROME	Q9w0g1 drosophila
27	41	100.0	586	2	Q9D5X3_MOUSE	Q9d5x3 mus musculu
28	41	100.0	601	1	KCNA5_MUSPF	P79197 mustela put
29	41	100.0	613	1	KCNA5_HUMAN	P22460 homo sapien
30	41	100.0	613	2	Q4VAJ1_HUMAN	Q4vaj1 homo sapien
31	41	100.0	631	2	Q4X0H0_ASPFU	Q4x0h0 aspergillus
32	41	100.0	631	2	Q5XIL5_RAT	Q5xil5 rattus norv
33	41	100.0	658	2	Q66HC8_RAT	Q66hc8 rattus norv
34	41	100.0	673	2	Q5EBP4_MOUSE	Q5ebp4 mus musculu
35	41	100.0	675	2	Q80WJ1_MOUSE	Q80wj1 mus musculu
36	41	100.0	753	2	Q6PIJ3_HUMAN	Q6pij3 homo sapien
37	41	100.0	760	2	Q8NBU7_HUMAN	Q8nbu7 homo sapien
38	41	100.0	775	2	Q4SME7_TETNG	Q4sme7 tetraodon n
39	41	100.0	780	2	Q9LK78_ARATH	Q9lk78 arabidopsis
40	41	100.0	781	2	Q823Y3_CHLCV	Q823y3 chlamydophi
41	41	100.0	782	2	Q8GX37_ARATH	Q8gx37 arabidopsis
42	41	100.0	812	2	Q9DWG7_RCMVM	Q9dwg7 rat cytomeg
43	41	100.0	824	2	Q5KQN9_ORYSA	Q5kqn9 oryza sativ
44	41	100.0	847	2	Q6UXZ5_HUMAN	Q6uxz5 homo sapien
45	41	100.0	847	2	Q5R9H3_PONPY	Q5r9h3 pongo pygma

ALIGNMENTS

RESULT 1
Q62KF6_BURMA
ID Q62KF6_BURMA PRELIMINARY; PRT; 40 AA.
AC Q62KF6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BMA1127;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidsen T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammoud Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000010; AAU47400.1; -; Genomic_DNA.
DR TIGR; BMA1127; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 40 AA; 4567 MW; 9DB8D75BCE486570 CRC64;

Query Match 100.0%; Score 41; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPPLP 7
              |||||||
Db         12 RPLPPLP 18
```

RESULT 2
Q8QS70_9BETA
ID Q8QS70_9BETA PRELIMINARY; PRT; 101 AA.
AC Q8QS70;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE UL17.
OS Pongine herpesvirus 4 (Chimpanzee cytomegalovirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=188763;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```
RX   MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;
RA   Davison A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
RA   Alcendor D.J., McGeoch D.J., Hayward G.S.;
RT   "The human cytomegalovirus genome revisited: comparison with the
RT   chimpanzee cytomegalovirus genome.";
RL   J. Gen. Virol. 84:17-28(2003).
DR   EMBL; AF480884; AAM00667.1; -; Genomic_DNA.
SQ   SEQUENCE   101 AA;  11945 MW;  366506511C6EA442 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 101;
  Best Local Similarity   100.0%;  Pred. No. 75;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db         42 RPLPPLP 48


RESULT 3
Q8STT3_ENCCU
ID   Q8STT3_ENCCU PRELIMINARY;      PRT;   128 AA.
AC   Q8STT3;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein ECU09_0750.
GN   OrderedLocusNames=ECU09_0750;
OS   Encephalitozoon cuniculi.
OC   Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX   NCBI_TaxID=6035;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=GB-M1;
RX   MEDLINE=21576510; PubMed=11719806; DOI=10.1038/35106579;
RA   Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA   Prensier G., Barbe V., Peyretaillade E., Brottier P., Wincker P.,
RA   Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA   Weissenbach J., Vivares C.P.;
RT   "Genome sequence and gene compaction of the eukaryote parasite
RT   Encephalitozoon cuniculi.";
RL   Nature 414:450-453(2001).
DR   EMBL; AL590451; CAD27048.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   128 AA;  14388 MW;  9096523574791EFC CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 128;
  Best Local Similarity   100.0%;  Pred. No. 97;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db          8 RPLPPLP 14


RESULT 4
Q67TQ3_ORYSA
ID   Q67TQ3_ORYSA PRELIMINARY;      PRT;   168 AA.
AC   Q67TQ3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Gallus gallus mRNA for attachment region binding protein (ARBP)-
DE   like.
GN   Name=B1342C04.28; Synonyms=P0025H07.47;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, BAC
RT   clone:B1342C04.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC
RT   clone:P0025H07.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP006057; BAD38468.1; -; Genomic_DNA.
DR   EMBL; AP005655; BAD38295.1; -; Genomic_DNA.
DR   Gramene; Q67TQ3; -.
SQ   SEQUENCE   168 AA;  17391 MW;  C70202755556154B CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db         96 RPLPPLP 102


RESULT 5
Q7XR59_ORYSA
ID   Q7XR59_ORYSA PRELIMINARY;      PRT;   199 AA.
AC   Q7XR59;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   OSJNBa0043A12.26 protein.
GN   Name=OSJNBa0043A12.26;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=12447439; DOI=10.1038/nature01183;
RA   Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA   Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA   Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA   Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA   Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA   Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA   Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA   Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA   Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA   Han B.;
RT   "Sequence and analysis of rice chromosome 4.";
RL   Nature 420:316-320(2002).
DR   EMBL; AL606619; CAE02821.2; -; Genomic_DNA.
DR   Gramene; Q7XR59; -.
DR   GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   199 AA;  22079 MW;  BFBB21C409444AEE CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 199;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db         78 RPLPPLP 84


RESULT 6
Q80WJ0_MOUSE
ID   Q80WJ0_MOUSE PRELIMINARY;      PRT;   271 AA.
AC   Q80WJ0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
```

```
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Gametogenetin protein 2.
GN   Name=Ggn;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=BALB/c;
RX   MEDLINE=22602683; PubMed=12574169; DOI=10.1074/jbc.M211023200;
RA   Lu B., Bishop C.E.;
RT   "Mouse GGN1 and GGN3, two germ cell-specific proteins from the single
RT   gene Ggn, interact with mouse POG and play a role in
RT   spermatogenesis.";
RL   J. Biol. Chem. 278:16289-16296(2003).
DR   EMBL; AF538033; AAP31498.1; -; mRNA.
DR   Ensembl; ENSMUSG00000031493; Mus musculus.
DR   MGI; MGI:2181461; Ggn.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005635; C:nuclear membrane; IDA.
DR   GO; GO:0005730; C:nucleolus; IDA.
DR   GO; GO:0046983; F:protein dimerization activity; IDA.
DR   GO; GO:0008104; P:protein localization; IDA.
SQ   SEQUENCE   271 AA;  28754 MW;  28A25610172AB42A CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        144 RPLPPLP 150


RESULT 7
Q6N908_RHOPA
ID   Q6N908_RHOPA PRELIMINARY;      PRT;   279 AA.
AC   Q6N908;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein precursor.
GN   OrderedLocusNames=RPA1743;
OS   Rhodopseudomonas palustris.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Rhodopseudomonas.
OX   NCBI_TaxID=1076;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CGA009 / ATCC BAA-98;
RX   PubMed=14704707; DOI=10.1038/nbt923;
RA   Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA   Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA   Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA   Harrison F.H., Gibson J., Harwood C.S.;
RT   "Complete genome sequence of the metabolically versatile
RT   photosynthetic bacterium Rhodopseudomonas palustris.";
RL   Nat. Biotechnol. 22:55-61(2004).
DR   EMBL; BX572598; CAE27184.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein; Signal.
FT   SIGNAL       1     24       Potential.
SQ   SEQUENCE   279 AA;  30610 MW;  22826BFA24CF75C9 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 279;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        227 RPLPPLP 233


RESULT 8
TNFL6_PIG
ID   TNFL6_PIG      STANDARD;      PRT;   282 AA.
AC   Q9BEA8; Q95M04; Q95N10;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)
DE   [Contains: Tumor necrosis factor ligand superfamily member 6, membrane
DE   form; Tumor necrosis factor ligand superfamily member 6, soluble
DE   form].
GN   Name=TNFSF6; Synonyms=FASL;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC   Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21322533; PubMed=11429161; DOI=10.1089/107999001300177493;
RA   Muneta Y., Shimoji Y., Inumaru S., Mori Y.;
RT   "Molecular cloning, characterization, and expression of porcine Fas
RT   ligand (CD95 ligand).";
RL   J. Interferon Cytokine Res. 21:305-312(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Guanxi bama miniature pig;
RA   Zhu N., Young Y.;
RT   "Molecular cloning and characterization of porcine Fas ligand cDNA.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Lymphoid;
RA   Tsuyuki S., Kono M., Bloom E.T.;
RT   "Cloning and potential utility of porcine Fas ligand: overexpression
RT   in porcine cells protects them from attack by human cytolytic cells.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Landrace x Large Yorkshire white; TISSUE=Thymocyte;
RX   MEDLINE=21653191; PubMed=11792426; DOI=10.1016/S0161-5890(01)00098-0;
RA   Motegi-Ishiyama Y., Nakajima Y., Hoka S., Takagaki Y.;
RT   "Porcine Fas-ligand gene: genomic sequence analysis and comparison
RT   with human gene.";
RL   Mol. Immunol. 38:581-586(2002).
CC   -!- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that
CC       transduces the apoptotic signal into cells. May be involved in
CC       cytotoxic T cell mediated apoptosis and in T cell development.
CC       TNFRSF6/FAS-mediated apoptosis may have a role in the induction of
CC       peripheral tolerance, in the antigen-stimulated suicide of mature
CC       T cells, or both. Binding to the decoy receptor TNFRSF6B/DcR3
CC       modulates its effects (By similarity).
CC   -!- SUBUNIT: Homotrimer (Probable).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC       similarity).
CC   -!- INDUCTION: By IL-18.
CC   -!- PTM: The soluble form derives from the membrane form by
CC       proteolytic processing (By similarity).
CC   -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AB027297; BAB40919.1; -; mRNA.
DR   EMBL; AY033634; AAK56449.1; -; mRNA.
DR   EMBL; AF397407; AAK84408.1; -; mRNA.
DR   EMBL; AB069764; BAB64291.1; -; Genomic_DNA.
DR   HSSP; P01375; 4TSV.
```

```
DR   InterPro; IPR008064; Fas_ligand.
DR   InterPro; IPR006053; TNF_abc.
DR   InterPro; IPR002961; TNF_C.
DR   InterPro; IPR006052; TNF_family.
DR   InterPro; IPR003636; TNF_subf.
DR   PANTHER; PTHR15161; Fas_ligand; 1.
DR   Pfam; PF00229; TNF; 1.
DR   PRINTS; PR01681; FASLIGAND.
DR   PRINTS; PR01234; TNECROSISFCT.
DR   PRINTS; PR01237; TNFC.
DR   ProDom; PD002012; TNF_subf; 1.
DR   SMART; SM00207; TNF; 1.
DR   PROSITE; PS00251; TNF_1; 1.
DR   PROSITE; PS50049; TNF_2; 1.
KW   Apoptosis; Cytokine; Glycoprotein; Signal-anchor; Transmembrane.
FT   CHAIN         1    282       Tumor necrosis factor ligand superfamily
FT                                member 6, membrane form.
FT   CHAIN       131    282       Tumor necrosis factor ligand superfamily
FT                                member 6, soluble form (By similarity).
FT   TOPO_DOM      1     82       Cytoplasmic (Potential).
FT   TRANSMEM     83    103       Signal-anchor for type II membrane
FT                                protein (Potential).
FT   TOPO_DOM    104    282       Extracellular (Potential).
FT   COMPBIAS      4     70       Pro-rich.
FT   COMPBIAS     45     56       Poly-Pro.
FT   SITE        130    131       Cleavage (By similarity).
FT   CARBOHYD    185    185       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    251    251       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    261    261       N-linked (GlcNAc. . .) (Potential).
FT   DISULFID    203    234       Potential.
FT   CONFLICT      5      5       F -> L (in Ref. 4).
FT   CONFLICT     57     57       T -> P (in Ref. 2).
SQ   SEQUENCE   282 AA;  31756 MW;  6743DAA1145671FB CRC64;

  Query Match             100.0%;  Score 41;  DB 1;  Length 282;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db         62 RPLPPLP 68


RESULT 9
Q9HCN7_HUMAN
ID   Q9HCN7_HUMAN PRELIMINARY;      PRT;   321 AA.
AC   Q9HCN7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Platelet glycoprotein VI-2.
GN   Name=GPVI;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20483673; PubMed=11027634; DOI=10.1006/bbrc.2000.3624;
RA   Ezumi Y., Uchiyama T., Takayama H.;
RT   "Molecular cloning, genomic structure, chromosomal localization, and
RT   alternative splice forms of the platelet collagen receptor
RT   glycoprotein VI.";
RL   Biochem. Biophys. Res. Commun. 277:27-36(2000).
DR   EMBL; AB043820; BAB12246.1; -; mRNA.
DR   HSSP; Q8NHL6; 1G0X.
DR   Ensembl; ENSG00000088053; Homo sapiens.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00409; IG; 2.
KW   Immunoglobulin domain.
SQ   SEQUENCE   321 AA;  35158 MW;  93FBF88945958345 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        289 RPLPPLP 295


RESULT 10
Q9UIF2_HUMAN
ID   Q9UIF2_HUMAN PRELIMINARY;      PRT;   339 AA.
AC   Q9UIF2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Platelet glycoprotein VI precursor (Platelet glycoprotein VI-1).
GN   Name=gpVI; Synonyms=GPVI;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Miura Y.;
RT   "Platelet glycoprotein VI.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20483673; PubMed=11027634; DOI=10.1006/bbrc.2000.3624;
RA   Ezumi Y., Uchiyama T., Takayama H.;
RT   "Molecular cloning, genomic structure, chromosomal localization, and
RT   alternative splice forms of the platelet collagen receptor
RT   glycoprotein VI.";
RL   Biochem. Biophys. Res. Commun. 277:27-36(2000).
CC   -!- INTERACTION:
CC       P30273:FCER1G; NbExp=1; IntAct=EBI-515278, EBI-515289;
CC       P06241:FYN; NbExp=1; IntAct=EBI-515278, EBI-515315;
CC       P07948:LYN; NbExp=2; IntAct=EBI-515278, EBI-79452;
CC       P07947:YES1; NbExp=1; IntAct=EBI-515278, EBI-515331;
DR   EMBL; AB035073; BAA89353.1; -; mRNA.
DR   EMBL; AB043819; BAB12245.1; -; mRNA.
DR   PIR; JC7509; JC7509.
DR   HSSP; Q8NHL6; 1G0X.
DR   IntAct; Q9UIF2; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005518; F:collagen binding; TAS.
DR   GO; GO:0004888; F:transmembrane receptor activity; TAS.
DR   GO; GO:0007167; P:enzyme linked receptor protein signaling pa. . .; TAS.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00409; IG; 2.
KW   Immunoglobulin domain; Signal.
FT   SIGNAL        1     20       Potential.
FT   CHAIN        21    339       Potential.
SQ   SEQUENCE   339 AA;  36923 MW;  4237576E95E030CC CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 339;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        307 RPLPPLP 313


RESULT 11
Q5YME9_NOCFA
ID   Q5YME9_NOCFA PRELIMINARY;      PRT;   378 AA.
```

```
AC   Q5YME9;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=pnf11170;
OS   Nocardia farcinica.
OG   Plasmid pNF1.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Nocardiaceae; Nocardia.
OX   NCBI_TaxID=37329;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=IFM 10152;
RX   PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA   Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA   Shiba T., Hattori M.;
RT   "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR   EMBL; AP006619; BAD60642.1; -; Genomic_DNA.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR001354; MR_MLE.
DR   Pfam; PF01188; MR_MLE; 1.
KW   Complete proteome; Hypothetical protein; Plasmid.
SQ   SEQUENCE   378 AA;  39948 MW;  251B906460428B4F CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 378;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db         12 RPLPPLP 18


RESULT 12
Q6PA50_XENLA
ID   Q6PA50_XENLA PRELIMINARY;      PRT;   401 AA.
AC   Q6PA50;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC68521 protein.
GN   Name=MGC68521;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RA   Klein S., Strausberg R.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC060454; AAH60454.1; -; mRNA.
SQ   SEQUENCE   401 AA;  44370 MW;  C88EBEE3F35F7333 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 401;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        144 RPLPPLP 150


RESULT 13
Q6DF84_XENLA
ID   Q6DF84_XENLA PRELIMINARY;      PRT;   404 AA.
AC   Q6DF84;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   LOC432336 protein.
GN   Name=LOC432336;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Oocytes;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Oocytes;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
```

```
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Oocytes;
RA   Klein S., Strausberg R.;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC076858; AAH76858.1; -; mRNA.
SQ   SEQUENCE   404 AA;  44557 MW;  49711922BDC71BC0 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 404;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        147 RPLPPLP 153


RESULT 14
Q872X2_NEUCR
ID   Q872X2_NEUCR PRELIMINARY;      PRT;   423 AA.
AC   Q872X2;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein B23B10.260.
GN   Name=B23B10.260;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   German Neurospora genome project;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BX284752; CAD70439.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   423 AA;  48356 MW;  BAEB36BEA38D1CE0 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 423;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        201 RPLPPLP 207


RESULT 15
Q8IRH4_DROME
ID   Q8IRH4_DROME PRELIMINARY;      PRT;   431 AA.
AC   Q8IRH4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CG9181-PD, isoform D.
GN   Name=Ptp61F; ORFNames=CG9181;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
```

```
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003471; AAN11476.1; -; Genomic_DNA.
DR   HSSP; P18031; 1ONY.
DR   Ensembl; CG9181; Drosophila melanogaster.
DR   FlyBase; FBgn0003138; CG9181.
DR   FlyBase; FBgn0003138; Ptp61F.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; IDA.
DR   GO; GO:0007411; P:axon guidance; IPI.
DR   GO; GO:0006470; P:protein amino acid dephosphorylation; IDA.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   InterPro; IPR000242; Tyr_PP.
DR   Pfam; PF00102; Y_phosphatase; 1.
DR   PRINTS; PR00700; PRTYPHPHTASE.
DR   SMART; SM00194; PTPc; 1.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW   Hydrolase.
SQ   SEQUENCE   431 AA;  48516 MW;  3E584E3881A82142 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 431;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPPLP 7
              |||||||
Db        274 RPLPPLP 280


Search completed: April  4, 2006, 13:15:20
Job time : 7.35079 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 3.47251 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-286
Perfect score: 41
Sequence: 1 RPLPIPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41	100.0	7	3	AAB17230	Aab17230 SH3 antag
2	41	100.0	7	3	AAB17231	Aab17231 SH3 antag
3	41	100.0	7	5	ABB73223	Abb73223 Src homol
4	41	100.0	7	5	ABB73224	Abb73224 Src homol
5	41	100.0	7	7	ADJ73377	Adj73377 SH3 antag
6	41	100.0	7	7	ADJ73378	Adj73378 SH3 antag
7	41	100.0	7	8	ADJ53012	Adj53012 CH1 delet
8	41	100.0	7	8	ADJ53011	Adj53011 CH1 delet
9	41	100.0	7	8	ADJ51973	Adj51973 CH1 delet
10	41	100.0	7	8	ADJ51972	Adj51972 CH1 delet
11	41	100.0	13	2	AAW05415	Aaw05415 Src SH3 d
12	41	100.0	13	2	AAW05482	Aaw05482 SH3-bindi
13	41	100.0	13	2	AAW16949	Aaw16949 Src SH3 d
14	41	100.0	13	2	AAW11101	Aaw11101 Src SH3 d
15	41	100.0	13	2	AAW25514	Aaw25514 SH3 synth
16	41	100.0	17	2	AAW05421	Aaw05421 SH3 domai
17	41	100.0	30	2	AAW16925	Aaw16925 Random re
18	41	100.0	30	2	AAW25488	Aaw25488 Random pe
19	41	100.0	30	2	AAW25489	Aaw25489 Random pe
20	41	100.0	237	4	AAB60970	Aab60970 E coli cy
21	40	97.6	7	2	AAW05076	Aaw05076 Proline-r
22	40	97.6	12	2	AAW05075	Aaw05075 Proline-r
23	40	97.6	15	2	AAW05067	Aaw05067 Proline-r
24	40	97.6	15	2	AAW38910	Aaw38910 Peptide r
25	40	97.6	579	7	ABM87190	Abm87190 Rice abio
26	39	95.1	10	2	AAR93548	Aar93548 Random 10
27	39	95.1	10	2	AAR93547	Aar93547 Random 10
28	39	95.1	14	2	AAR93467	Aar93467 GST-SRC p
29	39	95.1	15	2	AAW05068	Aaw05068 Proline-r
30	39	95.1	16	2	AAW25378	Aaw25378 Src SH3 d
31	39	95.1	20	2	AAW17009	Aaw17009 SRC SH3 d
32	39	95.1	79	8	ADK98541	Adk98541 Human imm
33	39	95.1	99	8	ADH22377	Adh22377 Human rec
34	39	95.1	248	6	ABP70948	Abp70948 Interleuk
35	39	95.1	277	4	AAO09572	Aao09572 Human pol
36	39	95.1	416	5	AAU77542	Aau77542 Bovine re
37	39	95.1	416	5	AAE23218	Aae23218 Cow recep
38	39	95.1	416	7	AAE39509	Aae39509 Bovine RA
39	39	95.1	416	7	ADG32003	Adg32003 Cow lung-
40	39	95.1	448	8	ADU83369	Adu83369 Human rec
41	39	95.1	553	2	AAW79159	Aaw79159 Zcytor7 c
42	39	95.1	553	4	ABB11582	Abb11582 Human cla
43	39	95.1	553	4	AAB85268	Aab85268 Human IL-
44	39	95.1	553	4	AAU04058	Aau04058 Human int
45	39	95.1	553	5	ABG67199	Abg67199 Human Int

ALIGNMENTS

RESULT 1
AAB17230
ID AAB17230 standard; peptide; 7 AA.
XX
AC AAB17230;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:286.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 297; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPIPP 7
|||||||
Db 1 RPLPIPP 7

RESULT 2
AAB17231
ID AAB17231 standard; peptide; 7 AA.
XX
AC AAB17231;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:287.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 297; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPIPP 7
|||||||
Db 1 RPLPIPP 7

RESULT 3
ABB73223
ID ABB73223 standard; peptide; 7 AA.
XX
AC ABB73223;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:286.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated

CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPIPP 7
              |||||||
Db          1 RPLPIPP 7
```

RESULT 4
ABB73224
ID ABB73224 standard; peptide; 7 AA.
XX
AC ABB73224;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:287.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPIPP 7
              |||||||
Db          1 RPLPIPP 7
```

RESULT 5
ADJ73377
ID ADJ73377 standard; peptide; 7 AA.
XX
AC ADJ73377;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 831.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 831; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific

CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPIPP 7
|||||||
Db 1 RPLPIPP 7

RESULT 6
ADJ73378
ID ADJ73378 standard; peptide; 7 AA.
XX
AC ADJ73378;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 832.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 832; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPIPP 7
|||||||
Db 1 RPLPIPP 7

RESULT 7
ADJ53012
ID ADJ53012 standard; peptide; 7 AA.
XX
AC ADJ53012;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID832.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 832; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart

CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPIPP 7
|||||||
Db 1 RPLPIPP 7

RESULT 8
ADJ53011
ID ADJ53011 standard; peptide; 7 AA.
XX
AC ADJ53011;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID831.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 831; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPIPP 7
|||||||
Db 1 RPLPIPP 7

RESULT 9
ADJ51973
ID ADJ51973 standard; peptide; 7 AA.
XX
AC ADJ51973;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID832.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 832; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-

```
CC   modulator or cytokine-agonist. The methods and compositions of the
CC   present invention are useful for the diagnosis, prevention and/or
CC   treatment of diseases or conditions associated with aberrant expression
CC   or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC   cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC   endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC   obstetric, haematologic, immunological, allergic, infectious,
CC   musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC   pediatric, psychiatric, renal or pulmonary disorders. The present
CC   sequence is that of a peptide which may be used during the creation of a
CC   mimetibody of the invention.
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 41;  DB 8;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db          1 RPLPIPP 7


RESULT 10
ADJ51972
ID   ADJ51972 standard; peptide; 7 AA.
XX
AC   ADJ51972;
XX
DT   06-MAY-2004  (first entry)
XX
DE   CH1 deleted mimetibody-related peptide SeqID831.
XX
KW   CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW   dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW   gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW   antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW   ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW   TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW   dental disorder; oral disorder; dermatological disorder; ear disorder;
KW   nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW   gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW   obstetric disorder; haematologic disorder; immunological disorder;
KW   allergic disorder; infectious disorder; musculoskeletal disorder;
KW   oncological disorder; neurological disorder; nutritional disorder;
KW   ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW   renal disorder; pulmonary disorder.
XX
OS   Unidentified.
OS   Synthetic.
XX
PN   WO2004002424-A2.
XX
PD   08-JAN-2004.
XX
PF   30-JUN-2003; 2003WO-US020495.
XX
PR   28-JUN-2002; 2002US-0392431P.
PR   19-SEP-2002; 2002US-0412144P.
XX
PA   (CENZ ) CENTOCOR INC.
XX
PI   Heavner GA,  Knight DM,  Ghrayeb J,  Scallon BJ,  Nesspor TC;
PI   Kutoloski KA;
XX
DR   WPI; 2004-082872/08.
XX
PT   New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT   diagnosing, preventing or treating cardiovascular, dermatologic,
PT   endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT   nutritional disorders.
XX
PS   Claim 15; SEQ ID NO 831; 123pp; English.
XX
CC   This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC   which encode them), compositions, methods and uses. The invention may be
CC   useful for the development of compounds with an osteopathic,
CC   cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC   gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC   immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC   antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC   respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC   modulator or cytokine-agonist. The methods and compositions of the
CC   present invention are useful for the diagnosis, prevention and/or
CC   treatment of diseases or conditions associated with aberrant expression
CC   or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC   cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC   endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC   obstetric, haematologic, immunological, allergic, infectious,
CC   musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC   pediatric, psychiatric, renal or pulmonary disorders. The present
CC   sequence is that of a peptide which may be used during the creation of a
CC   mimetibody of the invention.
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 41;  DB 8;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db          1 RPLPIPP 7


RESULT 11
AAW05415
ID   AAW05415 standard; peptide; 13 AA.
XX
AC   AAW05415;
XX
DT   24-FEB-1998  (first entry)
XX
DE   Src SH3 domain-binding peptide, T12SRC.4.
XX
KW   Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
KW   cellular signalling element; cellular structural element; malignancy;
KW   protein identification; functional domain; protein screening;
KW   cellular signal transduction process; binding peptide.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Biotin labelled"
FT   Modified-site   13
FT                   /note= "C-terminal amide"
XX
PN   WO9631625-A1.
XX
PD   10-OCT-1996.
XX
PF   04-APR-1996;   96WO-US004454.
XX
PR   07-APR-1995;   95US-00417872.
PR   03-APR-1996;   96US-00630915.
XX
PA   (CYTO-) CYTOGEN CORP.
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Hoffman N,  Kay BK,  Fowlkes DM,  Mcconnell SJ;
XX
DR   WPI; 1996-465045/46.
XX
PT   Identifying polypeptide(s) having specific functional domain (esp. SH3
```

PT domain) - comprises detecting selective binding to recognition unit,
PT regardless of sequence homology.
XX
PS Example; Page 81; 174pp; English.
XX
CC AAW05414 and AAW05415 represent variants of the Src Src-homology region 3
CC (SH3) domain-binding peptide termed pSrcCII (see AAW05412). These
CC sequences were used to probe human cDNA libraries to identify human SH3
CC domain containing proteins (such as AAW05400), that can be used in the
CC method of the invention. The method of the invention is for identifying
CC polypeptides containing functional domains of interest (especially SH3
CC domains). It comprises contacting a multivalent recognition unit (RU)
CC complex with a number of peptides and identifying polypeptides having a
CC selective binding affinity for the RU complex. The method is based on
CC functional similarities and does not rely on sequence similarities. Prior
CC methods only gave limited success for identifying proteins containing an
CC SH3 domain due to the minimal sequence homology among known SH3 proteins.
CC Multivalent RU complexes are particularly suited to screening for
CC polypeptides containing functional domains that are similar to, but not
CC identical in sequence to, the original target functional domain. The new
CC method enables proteins having a common function to be identified.
CC Identification of novel SH3 proteins will be useful for a better
CC understanding of cell growth, malignancy, signal transduction processes,
CC etc. New candidate drugs can be identified, and their specificities (e.g.
CC pharmacological activities) can be assessed using the method of the
CC invention
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 41; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPIPP 7
              |||||||
Db          4 RPLPIPP 10
```

RESULT 12
AAW05482
ID AAW05482 standard; peptide; 13 AA.
XX
AC AAW05482;
XX
DT 24-FEB-1998 (first entry)
XX
DE SH3-binding peptide T12SRC4.
XX
KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
KW cellular signalling element; cellular structural element; malignancy;
KW protein identification; functional domain; protein screening;
KW cellular signal transduction process; binding peptide.
XX
OS Synthetic.
XX
PN WO9631625-A1.
XX
PD 10-OCT-1996.
XX
PF 04-APR-1996; 96WO-US004454.
XX
PR 07-APR-1995; 95US-00417872.
PR 03-APR-1996; 96US-00630915.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Hoffman N, Kay BK, Fowlkes DM, Mcconnell SJ;
XX
DR WPI; 1996-465045/46.
XX
PT Identifying polypeptide(s) having specific functional domain (esp. SH3
PT domain) - comprises detecting selective binding to recognition unit,
PT regardless of sequence homology.
XX
PS Example; Fig 13; 174pp; English.
XX
CC AAW05445-W05492 represent Src-homology region 3 (SH2) domain binding
CC peptides. These sequences were used as parts of multivalent recognition
CC unit complexes used in the method of the invention. The method of the
CC invention is for identifying polypeptides containing functional domains
CC of interest (especially SH3 domains). It comprises contacting a
CC multivalent recognition unit (RU) complex with a number of peptides and
CC identifying polypeptides having a selective binding affinity for the RU
CC complex. The method is based on functional similarities and does not rely
CC on sequence similarities. Prior methods only gave limited success for
CC identifying proteins containing an SH3 domain due to the minimal sequence
CC homology among known SH3 proteins. Multivalent RU complexes are
CC particularly suited to screening for polypeptides containing functional
CC domains that are similar to, but not identical in sequence to, the
CC original target functional domain. The new method enables proteins having
CC a common function to be identified. Identification of novel SH3 proteins
CC will be useful for a better understanding of cell growth, malignancy,
CC signal transduction processes, etc. New candidate drugs can be
CC identified, and their specificities (e.g. pharmacological activities) can
CC be assessed using the method of the invention
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 41; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPIPP 7
              |||||||
Db          4 RPLPIPP 10
```

RESULT 13
AAW16949
ID AAW16949 standard; peptide; 13 AA.
XX
AC AAW16949;
XX
DT 27-JUN-1997 (first entry)
XX
DE Src SH3 domain-binding peptide used in signal transduction modulation.
XX
KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
OS Synthetic.
XX
PN WO9603649-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US009382.
XX
PR 22-JUL-1994; 94US-00278865.
PR 07-JUN-1995; 95US-00483555.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
XX
DR WPI; 1996-117151/12.
XX
PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.
XX
PS Disclosure; Page 24; 116pp; English.
XX
CC AAW16949-W16950 are peptides that bind to the Src SH3 domain. The SH3

```
CC   binding peptides are useful in modulating signal transduction pathways at
CC   the cellular level (especially protein tyrosine kinase-mediated),
CC   modulating oncogenic protein activity, or providing compounds for the
CC   development of drugs with the ability to modulate broad classes, as well
CC   as specific classes, of proteins involved in signal transduction and also
CC   for regulating the processing, trafficking or translation of RNA.
CC   Conjugates of the peptides with detectable labels or imaging agents are
CC   useful for imaging cells, tissues and organs in which Src or Src-related
CC   proteins are expressed
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db          4 RPLPIPP 10


RESULT 14
AAW11101
ID   AAW11101 standard; peptide; 13 AA.
XX
AC   AAW11101;
XX
DT   25-JUN-1997  (first entry)
XX
DE   Src SH3 domain-binding peptide used in signal transduction modulation.
XX
KW   Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW   protein tyrosine kinase; signal transduction; RNA processing;
KW   trafficking; translation.
XX
OS   Synthetic.
XX
PN   WO9603649-A1.
XX
PD   08-FEB-1996.
XX
PF   24-JUL-1995;   95WO-US009382.
XX
PR   22-JUL-1994;   94US-00278865.
PR   07-JUN-1995;   95US-00483555.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ;
XX
DR   WPI; 1996-117151/12.
XX
PT   Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT   proteins - useful for e.g. modulating signal transduction pathways at the
PT   cellular level, esp. protein tyrosine kinase-mediated.
XX
PS   Claim 35; Page 80; 116pp; English.
XX
CC   AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC   binding peptides are useful in modulating signal transduction pathways at
CC   the cellular level (especially protein tyrosine kinase-mediated),
CC   modulating oncogenic protein activity, or providing compounds for the
CC   development of drugs with the ability to modulate broad classes, as well
CC   as specific classes, of proteins involved in signal transduction and also
CC   for regulating the processing, trafficking or translation of RNA.
CC   Conjugates of the peptides with detectable labels or imaging agents are
CC   useful for imaging cells, tissues and organs in which Src or Src-related
CC   proteins are expressed
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db          4 RPLPIPP 10


RESULT 15
AAW25514
ID   AAW25514 standard; peptide; 13 AA.
XX
AC   AAW25514;
XX
DT   27-MAR-1998  (first entry)
XX
DE   SH3 synthetic binding peptide.
XX
KW   Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW   tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW   PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS   Synthetic.
XX
PN   WO9730074-A1.
XX
PD   21-AUG-1997.
XX
PF   14-FEB-1997;   97WO-US002298.
XX
PR   16-FEB-1996;   96US-00602999.
XX
PA   (CYTO-) CYTOGEN CORP.
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ,  Fowlkes DM;
PI   Rider JE;
XX
DR   WPI; 1997-424972/39.
XX
PT   Src homology region 3 binding peptide - used to activate Src tyrosine
PT   kinase(s) and to stimulate immune response by increasing production of
PT   certain lymphokine(s), e.g. interleukin-1.
XX
PS   Disclosure; Fig 7; 131pp; English.
XX
CC   The present sequence represents a Src SH3 synthetic binding peptide. SH3
CC   (Src homology region 3) binding peptides are selected from: (a) peptides
CC   which bind the SH3 domain of Cortactin; (b) peptides which bind the
CC   middle SH3 domain of Nck; (c) peptides which bind the SH3 domain of Abl;
CC   (d) peptides which bind the SH3 domain of Src; (e) peptides which bind
CC   the SH3 domain of PLC gamma; (f) peptides which bind the SH3 domain of
CC   p53bp2; (g) peptides which bind the amino-terminal SH3 domain of Crk; (h)
CC   peptides which bind the SH3 domain of Yes; and (i) peptides which bind
CC   the amino-terminal SH3 domain of Grb2. The purified binding peptides can
CC   be used in the method to identify inhibitors of their binding to their
CC   respective SH3 domains, which could be used to modulate the
CC   pharmacological activity of proteins or polypeptide containing the SH3
CC   domain. The peptides can also be used to activate Src or Src-related
CC   protein tyrosine kinases, to stimulate the immune response by increasing
CC   the production of certain lymphokines, e.g. tumour necrosis factor-alpha
CC   and interleukin-1, or to deliver a conjugated molecule to certain
CC   cellular compartments containing Src or Src related proteins
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 41;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db          4 RPLPIPP 10
```

Search completed: April 4, 2006, 13:07:43
Job time : 4.47251 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 1.14529 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-286
Perfect score: 41
Sequence: 1 RPLPIPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	97.6	222	2	C34223	transcription fact
2	40	97.6	284	2	B41224	homeotic protein p
3	39	95.1	203	2	S32799	hypothetical prote
4	39	95.1	416	1	A42879	advanced glycosyla
5	38	92.7	896	2	B43817	transforming prote
6	38	92.7	1267	2	T21340	hypothetical prote
7	38	92.7	1852	2	A37860	calcium channel pr
8	37	90.2	87	1	W4WL51	E4 protein - human
9	37	90.2	312	2	T16001	hypothetical prote
10	37	90.2	383	2	G86197	hypothetical prote
11	37	90.2	906	2	A43817	transforming prote
12	37	90.2	1103	2	JC5581	guanylate cyclase
13	37	90.2	1108	2	I59385	guanylate cyclase
14	37	90.2	1108	2	B55915	guanylate cyclase
15	37	90.2	1236	2	E70977	hypothetical prote
16	36	87.8	309	2	E84672	hypothetical prote
17	36	87.8	314	1	B40642	nodulation protein
18	36	87.8	361	2	AB2462	hypothetical prote
19	36	87.8	470	2	AD0888	SufI protein [impo
20	36	87.8	470	2	E91116	suppressor of ftsI
21	36	87.8	470	2	E85961	suppressor of ftsI
22	36	87.8	470	2	G65088	sufI protein precu
23	36	87.8	474	2	D75285	birA bifunctional
24	36	87.8	521	2	B84746	hypothetical prote
25	36	87.8	533	2	AC0414	probable exported
26	36	87.8	571	2	C75530	conserved hypothet
27	36	87.8	572	2	B48521	bilirubin oxidase
28	36	87.8	584	2	S76424	hypothetical prote
29	36	87.8	663	2	T40493	hnf-3/forkhead tra
30	36	87.8	701	2	S61239	hypothetical prote
31	36	87.8	711	2	S68443	double-stranded RN
32	36	87.8	800	2	T19627	hypothetical prote
33	36	87.8	1307	2	T17453	ERG-associated pro
34	35	85.4	113	2	T19198	hypothetical prote
35	35	85.4	132	2	T49589	hypothetical prote
36	35	85.4	156	2	D70909	probable two-compo
37	35	85.4	215	2	T41363	hypothetical prote
38	35	85.4	238	2	T40820	proline-rich prote
39	35	85.4	253	2	T08668	hypothetical prote
40	35	85.4	273	2	D98348	hypothetical prote
41	35	85.4	279	2	A53062	Fas ligand - mouse
42	35	85.4	344	2	S59043	spliling factor SR
43	35	85.4	455	2	G75473	probable carotenoi
44	35	85.4	645	2	T39614	kinase-binding pro
45	35	85.4	646	2	T48644	negative regulator

ALIGNMENTS

RESULT 1
C34223
transcription factor ATF-3 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: C34223
R;Hai, T.; Liu, F.; Coukos, W.J.; Green, M.R.
Genes Dev. 3, 2083-2090, 1989
A;Title: Transcription factor ATF cDNA clones: an extensive family of leucine zipper p
A;Reference number: A91622; MUID:90185187; PMID:2516827
A;Accession: C34223
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-222 <HA3>
A;Cross-references: UNIPROT:P18847; UNIPARC:UPI00001764AE
C;Genetics:
A;Gene: GDB:ATF3
A;Cross-references: GDB:370911
C;Superfamily: fos transforming protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; transcription regulation
F;122-162/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 97.6%; Score 40; DB 2; Length 222;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPIPP 7
||||:||
Db 61 RPLPVPP 67

RESULT 2
B41224
homeotic protein pMUR10F - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 31-Dec-2004
C;Accession: B41224
R;Kennedy, M.A.; Gonzalez-Sarmiento, R.; Kees, U.R.; Lampert, F.; Dear, N.; Boehm, T.;
Proc. Natl. Acad. Sci. U.S.A. 88, 8900-8904, 1991
A;Title: HOX11, a homeobox-containing T-cell oncogene on human chromosome 10q24.
A;Reference number: A41224; MUID:92020958; PMID:1681546
A;Accession: B41224
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-284 <KEN>
A;Cross-references: UNIPROT:Q61663; UNIPARC:UPI000002338B; GB:M75953; NID:g193843; PIDN
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;158-214/Domain: homeobox homology <HOX>

Query Match 97.6%; Score 40; DB 2; Length 284;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPIPP 7
              ||||:||
Db         87 RPLPVPP 93


RESULT 3
S32799
hypothetical protein 1 - Xanthomonas sp. transposon Tn5053 (fragment)
C;Species: Xanthomonas sp.
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 11-Jan-2000
C;Accession: S32799
R;Kholodii, G.Y.; Yurieva, O.V.; Lomovskaya, O.L.; Gorlenko, Z.M.; Mindlin, S.Z.; Nikifc
J. Mol. Biol. 230, 1103-1107, 1993
A;Title: Tn5053, a mercury resistance transposon with integron's ends.
A;Reference number: S32795; MUID:93253772; PMID:8387603
A;Accession: S32799
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-203 <KHO>
A;Cross-references: UNIPARC:UPI000017911D; EMBL:L03735; NID:g154911; PIDN:AAA91612.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1992
C;Genetics:
A;Mobile element: transposon Tn5053
C;Superfamily: Klebsiella transposase

  Query Match             95.1%;  Score 39;  DB 2;  Length 203;
  Best Local Similarity   85.7%;  Pred. No. 23;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||||:||
Db        105 RPLPLPP 111


RESULT 4
A42879
advanced glycosylation end-products receptor precursor - bovine
N;Alternate names: advanced glycosylation end product-binding protein, 35K; glycoprotein
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Mar-1993 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
C;Accession: A42879; A42878; S27949
R;Neeper, M.; Schmidt, A.M.; Brett, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.; St
J. Biol. Chem. 267, 14998-15004, 1992
A;Title: Cloning and expression of a cell surface receptor for advanced glycosylation en
A;Reference number: A42879; MUID:92340547; PMID:1378843
A;Accession: A42879
A;Molecule type: mRNA
A;Residues: 1-416 <NEE>
A;Cross-references: UNIPROT:Q28173; UNIPARC:UPI00001330EE; GB:M91212; NID:g163650; PIDN:
A;Experimental source: lung
A;Note: sequence extracted from NCBI backbone (NCBIP:109436)
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Schmidt, A.M.; Vianna, M.; Gerlach, M.; Brett, J.; Ryan, J.; Kao, J.; Esposito, C.; He
J. Biol. Chem. 267, 14987-14997, 1992
A;Title: Isolation and characterization of two binding proteins for advanced glycosylati
A;Reference number: A42878; MUID:92340546; PMID:1321822
A;Accession: A42878
A;Molecule type: protein
A;Residues: 23-24,'X',26-37,'X',39-49,'XX',52-54 <SCH>
A;Cross-references: UNIPARC:UPI00000876EC
A;Experimental source: endothelial cells
A;Note: sequence extracted from NCBI backbone (NCBIP:109434)
C;Comment: Advanced glycosylation end products are heterogeneous nonenzymatically glycos
cellular function, thus contributing to tissue lesions in diabetes.
C;Comment: This receptor appears also to mediate the effects of amyloid beta peptide on
ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.
C;Function:
A;Description: neuronal receptor for amphoterin, a DNA-binding protein involved in neuri
C;Superfamily: advanced glycosylation end products receptor; immunoglobulin homology
C;Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-416/Product: advanced glycosylation end-products receptor RAGE #status predicted <M
F;23-354/Domain: extracellular #status predicted <EXT>
F;31-100/Domain: immunoglobulin homology <IM1>
F;136-209/Domain: immunoglobulin homology <IM2>
F;262-313/Domain: immunoglobulin homology <IM3>
F;355-372/Domain: transmembrane #status predicted <TMM>
F;373-416/Domain: intracellular #status predicted <INT>
F;25,80/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;38-98,143-207,269-311/Disulfide bonds: #status predicted

  Query Match             95.1%;  Score 39;  DB 1;  Length 416;
  Best Local Similarity   85.7%;  Pred. No. 50;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||||:||
Db        286 RPLPLPP 292


RESULT 5
B43817
transforming protein (cbl) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 17-Mar-1999
C;Accession: B43817
R;Blake, T.J.; Shapiro, M.; Morse III, H.C.; Langdon, W.Y.
Oncogene 6, 653-657, 1991
A;Title: The sequences of the human and mouse c-cbl proto-oncogenes show v-cbl was gene
A;Reference number: A43817; MUID:91232862; PMID:2030914
A;Accession: B43817
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-896 <BLA>
A;Cross-references: UNIPARC:UPI000017C87A; EMBL:X57111

  Query Match             92.7%;  Score 38;  DB 2;  Length 896;
  Best Local Similarity   85.7%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              :||||||
Db        512 KPLPIPP 518


RESULT 6
T21340
hypothetical protein F45H11.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T21340; T22252
R;McMurray, A.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19409
A;Accession: T21340
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1267 <WIL>
A;Cross-references: UNIPROT:Q93564; UNIPARC:UPI0000083350; EMBL:Z78418; PIDN:CAB01699.1
A;Experimental source: clone F25D7
R;Kelly, P.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z19537
A;Accession: T22252
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1267 <WI2>
A;Cross-references: UNIPARC:UPI0000083350; EMBL:Z78420; PIDN:CAB01711.1; GSPDB:GN00019;
A;Experimental source: clone F45H11
C;Genetics:
A;Gene: CESP:F45H11.4
A;Map position: 1
A;Introns: 38/3; 90/2; 149/3; 207/1; 356/2; 413/2; 458/2; 520/3; 691/3; 777/2; 796/2; 8

  Query Match             92.7%;  Score 38;  DB 2;  Length 1267;
```

```
  Best Local Similarity   71.4%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||:|:||
Db       1244 RPIPVPP 1250


RESULT 7
A37860
calcium channel protein alpha-1 chain, skeletal muscle - common carp
C;Species: Cyprinus carpio (common carp)
C;Date: 31-May-1991 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: A37860
R;Grabner, M.; Friedrich, K.; Knaus, H.G.; Striessnig, J.; Scheffauer, F.; Staudinger, R
Proc. Natl. Acad. Sci. U.S.A. 88, 727-731, 1991
A;Title: Calcium channels from Cyprinus carpio skeletal muscle.
A;Reference number: A37860; MUID:91126068; PMID:1846962
A;Accession: A37860
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1852 <GRA>
A;Cross-references: UNIPROT:P22316; UNIPARC:UPI0000127281; GB:M62554; GB:M37203; NID:g21
C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C;Keywords: glycoprotein; phosphoprotein; skeletal muscle; transmembrane protein
F;74-90/Domain: transmembrane #status predicted <TR01>
F;108-131/Domain: transmembrane #status predicted <TR02>
F;140-158/Domain: transmembrane #status predicted <TR03>
F;212-234/Domain: transmembrane #status predicted <TR05>
F;328-350/Domain: transmembrane #status predicted <TR06>
F;448-466/Domain: transmembrane #status predicted <TR07>
F;484-501/Domain: transmembrane #status predicted <TR08>
F;514-530/Domain: transmembrane #status predicted <TR09>
F;577-596/Domain: transmembrane #status predicted <TR11>
F;650-676/Domain: transmembrane #status predicted <TR12>
F;817-834/Domain: transmembrane #status predicted <TR13>
F;853-870/Domain: transmembrane #status predicted <TR14>
F;884-901/Domain: transmembrane #status predicted <TR15>
F;947-966/Domain: transmembrane #status predicted <TR17>
F;1057-1084/Domain: transmembrane #status predicted <TR18>
F;1135-1153/Domain: transmembrane #status predicted <TR19>
F;1169-1188/Domain: transmembrane #status predicted <TR20>
F;1197-1215/Domain: transmembrane #status predicted <TR21>
F;1291-1310/Domain: transmembrane #status predicted <TR23>
F;1377-1402/Domain: transmembrane #status predicted <TR24>
F;99,102,274,470,813,1157,1269,1485,1703,1713,1745,1760,1848/Binding site: carbohydrate
F;407/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predic
F;1471,1523,1738/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #st

  Query Match             92.7%;  Score 38;  DB 2;  Length 1852;
  Best Local Similarity   71.4%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||:|:||
Db       1832 RPIPVPP 1838


RESULT 8
W4WL51
E4 protein - human papillomavirus type 51
C;Species: human papillomavirus type 51
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: C40415
R;Lungu, O.; Crum, C.P.; Silverstein, S.J.
J. Virol. 65, 4216-4225, 1991
A;Title: Biologic properties and nucleotide sequence analysis of human papillomavirus ty
A;Reference number: A40415; MUID:91303675; PMID:1649326
A;Accession: C40415
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-87 <LUN>
A;Cross-references: UNIPROT:P26548; UNIPARC:UPI0000138387; GB:M62877
C;Superfamily: papillomavirus E4 protein
C;Keywords: early protein

  Query Match             90.2%;  Score 37;  DB 1;  Length 87;
  Best Local Similarity   71.4%;  Pred. No. 20;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||:|:||
Db         24 RPIPLPP 30


RESULT 9
T16001
hypothetical protein F09E5.12 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16001
R;Chissoe, S.
submitted to the EMBL Data Library, September 1995
A;Description: The sequence of C. elegans cosmid F09E5.
A;Reference number: Z18444
A;Accession: T16001
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-312 <CHI>
A;Cross-references: UNIPROT:Q19261; UNIPARC:UPI00000826A8; EMBL:U37429; NID:g1019949;
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:F09E5.12
A;Introns: 30/3; 125/1; 162/3

  Query Match             90.2%;  Score 37;  DB 2;  Length 312;
  Best Local Similarity   71.4%;  Pred. No. 76;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||:|:||
Db        255 RPIPLPP 261


RESULT 10
G86197
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G86197
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86197
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-383 <STO>
A;Cross-references: UNIPROT:Q9LNC6; UNIPARC:UPI00000AC01F; GB:AE005172; NID:g8844126; E
C;Genetics:
A;Map position: 1

  Query Match             90.2%;  Score 37;  DB 2;  Length 383;
  Best Local Similarity   71.4%;  Pred. No. 94;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||:|:||
```

```
Db        191 RPVPVPP 197


RESULT 11
A43817
transforming protein (cbl) - human
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 09-Jul-2004
C;Accession: A43817
R;Blake, T.J.; Shapiro, M.; Morse III, H.C.; Langdon, W.Y.
Oncogene 6, 653-657, 1991
A;Title: The sequences of the human and mouse c-cbl proto-oncogenes show v-cbl was gener
A;Reference number: A43817; MUID:91232862; PMID:2030914
A;Accession: A43817
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-906 <BLA>
A;Cross-references: UNIPROT:P22681; UNIPARC:UPI00001271AA; EMBL:X57110; NID:g29730; PIDN
C;Keywords: DNA binding
F;377-425/Domain: RING finger homology <RRN>

  Query Match             90.2%;  Score 37;  DB 2;  Length 906;
  Best Local Similarity   71.4%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              :|||:||
Db        531 KPLPVPP 537


RESULT 12
JC5581
guanylate cyclase (EC 4.6.1.2) ROS-GC2 precursor - bovine
N;Alternate names: guanyl cyclase; guanylyl cyclase
C;Species: Bos primigenius taurus (cattle)
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 09-Jul-2004
C;Accession: JC5581
R;Goraczniak, R.; Duda, T.; Sharma, R.K.
Biochem. Biophys. Res. Commun. 234, 666-670, 1997
A;Title: Structural and functional characterization of a second subfamily member of the
A;Reference number: JC5581; MUID:97318835; PMID:9175772
A;Accession: JC5581
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1103 <GOR>
A;Cross-references: UNIPROT:O02740; UNIPARC:UPI0000128C1F; GB:U95958; NID:g2072999; PIDN
A;Experimental source: retina
C;Comment: This enzyme belongs to the subfamily of calcium-modulated rod outer segment m
C;Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain homo
C;Keywords: phosphorus-oxygen lyase
F;1-50/Domain: signal sequence #status predicted <SIG>
F;51-465/Domain: extracellular #status predicted <EXT>
F;466-490/Domain: transmembrane #status predicted <TRM>
F;523-816/Domain: protein kinase homology <KIN>
F;836-1064/Domain: guanylate cyclase catalytic domain homology <GCC>
F;1099-1103/Region: signature

  Query Match             90.2%;  Score 37;  DB 2;  Length 1103;
  Best Local Similarity   71.4%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              :|||:||
Db       1068 KPLPVPP 1074


RESULT 13
I59385
guanylate cyclase (EC 4.6.1.2) 2D-like precursor, retinal - human
C;Species: Homo sapiens (man)
C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 09-Jul-2004
C;Accession: I59385
R;Lowe, D.G.; Dizhoor, A.M.; Liu, K.; Gu, Q.; Spencer, M.; Laura, R.; Lu, L.; Hurley, J.
Proc. Natl. Acad. Sci. U.S.A. 92, 5535-5539, 1995
A;Title: Cloning and expression of a second photoreceptor-specific membrane retina guan
A;Reference number: I59385; MUID:95296345; PMID:7777544
A;Accession: I59385
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1108 <RES>
A;Cross-references: UNIPROT:P51841; UNIPARC:UPI0000128C20; GB:L37378; NID:g945224; PIDN
C;Genetics:
A;Gene: GDB:GUC2DL
A;Cross-references: GDB:701610
C;Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain hom
C;Keywords: cGMP biosynthesis; glycoprotein; phosphorus-oxygen lyase; transmembrane pro
F;523-816/Domain: protein kinase homology <KIN>
F;836-1064/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             90.2%;  Score 37;  DB 2;  Length 1108;
  Best Local Similarity   71.4%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              :|||:||
Db       1068 KPLPVPP 1074


RESULT 14
B55915
guanylate cyclase (EC 4.6.1.2) 2F precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C;Accession: B55915
R;Yang, R.B.; Foster, D.C.; Garbers, D.L.; Fuelle, H.J.
Proc. Natl. Acad. Sci. U.S.A. 92, 602-606, 1995
A;Title: Two membrane forms of guanylyl cyclase found in the eye.
A;Reference number: A55915; MUID:95132648; PMID:7831337
A;Accession: B55915
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1108 <YAN>
A;Cross-references: UNIPROT:P51842; UNIPARC:UPI0000128C21; GB:L36030; NID:g780358; PIDN
C;Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain hom
C;Keywords: cGMP biosynthesis; glycoprotein; phosphorus-oxygen lyase; transmembrane pro
F;523-816/Domain: protein kinase homology <KIN>
F;836-1064/Domain: guanylate cyclase catalytic domain homology <GCC>

  Query Match             90.2%;  Score 37;  DB 2;  Length 1108;
  Best Local Similarity   71.4%;  Pred. No. 2.9e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              :|||:||
Db       1068 KPLPVPP 1074


RESULT 15
E70977
hypothetical protein Rv3447c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70977
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70977
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1236 <COL>
A;Cross-references: UNIPROT:O06264; UNIPARC:UPI00000C150F; GB:Z95389; GB:AL123456; NID:
A;Experimental source: strain H37Rv
```

C;Genetics:
A;Gene: Rv3447c

Query Match 90.2%; Score 37; DB 2; Length 1236;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPIPP 7
              ||:|:||
Db       1207 RPMPLPP 1213
```

Search completed: April 4, 2006, 13:17:26
Job time : 2.14529 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 5.35079 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-286
Perfect score: 41
Sequence: 1 RPLPIPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	41	100.0	224	2	Q6EZE1_ECOLI	Q6eze1 escherichia
2	41	100.0	237	2	Q56UC1_ECOLI	Q56uc1 escherichia
3	41	100.0	237	2	Q47088_ECOLI	Q47088 escherichia
4	41	100.0	506	2	Q4ID33_GIBZE	Q4id33 gibberella
5	41	100.0	583	2	Q8FQU9_COREF	Q8fqu9 corynebacte
6	41	100.0	732	2	Q8J1Y5_ASHGO	Q8j1y5 ashbya goss
7	41	100.0	790	2	Q6NU24_XENLA	Q6nu24 xenopus lae
8	41	100.0	793	1	PHK_STRAW	Q82nm9 streptomyce
9	41	100.0	2275	2	Q8QS41_9BETA	Q8qs41 pongine her
10	40	97.6	175	2	Q5VPS2_ORYSA	Q5vps2 oryza sativ
11	40	97.6	284	1	TLX2_HUMAN	O43763 homo sapien
12	40	97.6	284	1	TLX2_MOUSE	Q61663 mus musculu
13	40	97.6	852	2	Q8QRS5_9BETA	Q8qrs5 pongine her
14	40	97.6	1074	2	Q4PG90_USTMA	Q4pg90 ustilago ma
15	39	95.1	88	2	Q993Z6_HPV26	Q993z6 human papil
16	39	95.1	88	2	Q9IR55_HPV26	Q9ir55 human papil
17	39	95.1	198	2	Q67VV2_ORYSA	Q67vv2 oryza sativ
18	39	95.1	209	2	Q96SH7_HUMAN	Q96sh7 homo sapien
19	39	95.1	278	2	Q73J56_TREDE	Q73j56 treponema d
20	39	95.1	416	1	RAGE_BOVIN	Q28173 bos taurus
21	39	95.1	552	2	Q6Z9X4_ORYSA	Q6z9x4 oryza sativ
22	39	95.1	553	1	I20RA_HUMAN	Q9uhf4 homo sapien
23	39	95.1	854	2	Q4QJ51_LEIMA	Q4qj51 leishmania
24	39	95.1	857	2	Q4PC21_USTMA	Q4pc21 ustilago ma
25	39	95.1	1104	2	Q8GUZ9_POPTM	Q8guz9 populus tre
26	39	95.1	3119	2	Q8IHM0_PLAF7	Q8ihm0 plasmodium
27	38	92.7	130	2	Q67R30_SYMTH	Q67r30 symbiobacte
28	38	92.7	287	2	Q7U7P1_SYNPX	Q7u7p1 synechococc
29	38	92.7	323	2	Q82J46_STRAW	Q82j46 streptomyce
30	38	92.7	440	2	Q568P5_BRARE	Q568p5 brachydanio
31	38	92.7	578	2	Q4P1B6_USTMA	Q4p1b6 ustilago ma
32	38	92.7	810	2	Q61DM6_CAEBR	Q61dm6 caenorhabdi
33	38	92.7	903	2	Q98TY6_CHICK	Q98ty6 gallus gall
34	38	92.7	913	1	CBL_MOUSE	P22682 mus musculu
35	38	92.7	1243	2	Q93564_CAEEL	Q93564 caenorhabdi
36	38	92.7	1641	2	Q9GRZ3_CAEEL	Q9grz3 caenorhabdi
37	38	92.7	1847	2	Q6RKB0_BRARE	Q6rkb0 brachydanio
38	38	92.7	1852	1	CAC1S_CYPCA	P22316 cyprinus ca
39	38	92.7	3855	2	Q4SYK6_TETNG	Q4syk6 tetraodon n
40	37	90.2	87	1	VE4_HPV51	P26548 human papil
41	37	90.2	117	2	Q5SIT2_THET8	Q5sit2 thermus the
42	37	90.2	117	2	Q72J63_THET2	Q72j63 thermus the
43	37	90.2	161	2	Q88C27_PSEPK	Q88c27 pseudomonas
44	37	90.2	176	2	Q5TP55_ANOGA	Q5tp55 anopheles g
45	37	90.2	312	2	Q19261_CAEEL	Q19261 caenorhabdi

ALIGNMENTS

RESULT 1
Q6EZE1_ECOLI
ID Q6EZE1_ECOLI PRELIMINARY; PRT; 224 AA.
AC Q6EZE1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MBU. E 412;
RX PubMed=15748977;
RA Bouzari S., Oloomi M., Oswald E.;
RT "Detection of the cytolethal distending toxin locus cdtB among
RT diarrheagenic Escherichia coli isolates from humans in Iran.";
RL Res. Microbiol. 156:137-144(2005).
DR EMBL; AF373206; AAT65834.2; -; Genomic_DNA.
DR GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003558; CDtoxinA.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF03498; CDtoxinA; 1.
DR PIRSF; PIRSF036516; CDT_A; 1.
DR PRINTS; PR01387; CDTOXINA.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 224 AA; 24123 MW; 347CE412AEB95961 CRC64;

Query Match 100.0%; Score 41; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPIPP 7
              |||||||
Db        208 RPLPIPP 214
```

RESULT 2
Q56UC1_ECOLI
ID Q56UC1_ECOLI PRELIMINARY; PRT; 237 AA.
AC Q56UC1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Cytolethal distending toxin type IV subunit A.
GN Name=cdtA;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=28C;
RA   Ledger N., Fujiwara T., Boury M., Sugai M., Oswald E.;
RT   "Escherichia coli 28C (O75) cytolethal distending toxin type IV.";
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY578329; AAT92047.1; -; Genomic_DNA.
SQ   SEQUENCE   237 AA;  25483 MW;  45C29A34455ECBDA CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 237;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db        221 RPLPIPP 227


RESULT 3
Q47088_ECOLI
ID   Q47088_ECOLI PRELIMINARY;      PRT;   237 AA.
AC   Q47088;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   CdtA.
GN   Name=cdtA;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=E6468/62;
RX   MEDLINE=94086109; PubMed=8262635;
RA   Scott D.A., Kaper J.B.;
RT   "Cloning and sequencing of the genes encoding Escherichia coli
RT   cytolethal distending toxin.";
RL   Infect. Immun. 62:244-251(1994).
DR   EMBL; U03293; AAD10621.1; -; Genomic_DNA.
DR   GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR   GO; GO:0005529; F:sugar binding; IEA.
DR   GO; GO:0009405; P:pathogenesis; IEA.
DR   InterPro; IPR003558; CDtoxinA.
DR   InterPro; IPR000772; Ricin_B_lectin.
DR   Pfam; PF03498; CDtoxinA; 1.
DR   PIRSF; PIRSF036516; CDT_A; 1.
DR   PRINTS; PR01387; CDTOXINA.
DR   PROSITE; PS50231; RICIN_B_LECTIN; 1.
SQ   SEQUENCE   237 AA;  25596 MW;  6D7EC323E4968E4E CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 237;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db        221 RPLPIPP 227


RESULT 4
Q4ID33_GIBZE
ID   Q4ID33_GIBZE PRELIMINARY;      PRT;   506 AA.
AC   Q4ID33;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG04875.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000198; EAA74203.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   506 AA;  56193 MW;  6DC63195B023F444 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 506;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db        413 RPLPIPP 419


RESULT 5
Q8FQU9_COREF
ID   Q8FQU9_COREF PRELIMINARY;      PRT;   583 AA.
AC   Q8FQU9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative oxidase.
GN   OrderedLocusNames=CE1018;
OS   Corynebacterium efficiens.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=152794;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX   MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA   Nishio Y., Nakamura Y., Kawarabayasi Y., Usuda Y., Kimura E.,
RA   Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA   Gojobori T.;
RT   "Comparative complete genome sequence analysis of the amino acid
RT   replacements responsible for the thermostability of Corynebacterium
RT   efficiens.";
RL   Genome Res. 13:1572-1579(2003).
DR   EMBL; BA000035; BAC17828.1; -; Genomic_DNA.
DR   HSSP; P36649; 1N68.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   InterPro; IPR011706; Cu-oxidase_2.
DR   InterPro; IPR011707; Cu-oxidase_3.
```

```
DR   InterPro; IPR002355; Cu_ox_copper_BS.
DR   InterPro; IPR006311; Tat.
DR   Pfam; PF07731; Cu-oxidase_2; 1.
DR   Pfam; PF07732; Cu-oxidase_3; 1.
DR   TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
DR   PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW   Complete proteome.
SQ   SEQUENCE   583 AA;  63559 MW;  AE5736B1AF2BECCF CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 583;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db        109 RPLPIPP 115


RESULT 6
Q8J1Y5_ASHGO
ID   Q8J1Y5_ASHGO PRELIMINARY;      PRT;   732 AA.
AC   Q8J1Y5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Wal1 protein.
GN   Name=WAL1;
OS   Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX   NCBI_TaxID=33169;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15367585; DOI=10.1242/jcs.01377;
RA   Walther A., Wendland J.;
RT   "Apical localization of actin patches and vacuolar dynamics in Ashbya
RT   gossypii depend on the WASP homolog Wallp.";
RL   J. Cell Sci. 117:4947-4958(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Wendland J.W., Walther A., Philippsen P.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY144115; AAN28957.1; -; Genomic_DNA.
DR   HSSP; O08816; 1MKE.
DR   InterPro; IPR000697; EVH1.
DR   InterPro; IPR011993; PH_type.
DR   InterPro; IPR001960; WH1.
DR   InterPro; IPR003124; WH2_actin_bd.
DR   Pfam; PF00568; WH1; 1.
DR   Pfam; PF02205; WH2; 1.
DR   SMART; SM00461; WH1; 1.
SQ   SEQUENCE   732 AA;  77695 MW;  438CF77D357E849D CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 732;
  Best Local Similarity   100.0%;  Pred. No. 4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db        312 RPLPIPP 318


RESULT 7
Q6NU24_XENLA
ID   Q6NU24_XENLA PRELIMINARY;      PRT;   790 AA.
AC   Q6NU24;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   MGC81305 protein.
GN   Name=MGC81305;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR   EMBL; BC068778; AAH68778.1; -; mRNA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR   GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR   GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR008271; Ser_thr_pkin_AS.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   InterPro; IPR001245; Tyr_pkinase.
DR   Pfam; PF00069; Pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   SMART; SM00220; S_TKc; 1.
DR   SMART; SM00219; TyrKc; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW   ATP-binding; Cell cycle; Cell division; Kinase; Nucleotide-binding;
KW   Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE   790 AA;  88982 MW;  53450731EF8F0FE1 CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 790;
  Best Local Similarity   100.0%;  Pred. No. 4.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db        525 RPLPIPP 531


RESULT 8
```

```
PHK_STRAW
ID   PHK_STRAW      STANDARD;      PRT;   793 AA.
AC   Q82NM9;
DT   10-MAY-2005 (Rel. 47, Created)
DT   10-MAY-2005 (Rel. 47, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Probable phosphoketolase (EC 4.1.2.-).
GN   OrderedLocusNames=SAV1273;
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
CC   -!- COFACTOR: Thiamine pyrophosphate (Potential).
CC   -!- SIMILARITY: Belongs to the XFP family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; BA000030; BAC68983.1; -; Genomic_DNA.
DR   HAMAP; MF_01403; -; 1.
DR   InterPro; IPR012109; Phosphoketolase.
DR   InterPro; IPR000399; TPP_bd.
DR   InterPro; IPR005593; XFP.
DR   Pfam; PF03894; XFP; 1.
DR   PIRSF; PIRSF017245; Phosphoketolase; 1.
DR   PROSITE; PS60002; PHOSPHOKETOLASE_1; 1.
DR   PROSITE; PS60003; PHOSPHOKETOLASE_2; 1.
DR   PROSITE; PS00187; TPP_ENZYMES; FALSE_NEG.
KW   Complete proteome; Flavoprotein; Lyase; Thiamine pyrophosphate.
SQ   SEQUENCE   793 AA;  88062 MW;  39D02CF9AF57E783 CRC64;

  Query Match             100.0%;  Score 41;  DB 1;  Length 793;
  Best Local Similarity   100.0%;  Pred. No. 4.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db        376 RPLPIPP 382


RESULT 9
Q8QS41_9BETA
ID   Q8QS41_9BETA PRELIMINARY;      PRT;  2275 AA.
AC   Q8QS41;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Tegument protein UL48.
OS   Pongine herpesvirus 4 (Chimpanzee cytomegalovirus).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
OX   NCBI_TaxID=188763;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;
RA   Davison A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
RA   Alcendor D.J., McGeoch D.J., Hayward G.S.;
RT   "The human cytomegalovirus genome revisited: comparison with the
RT   chimpanzee cytomegalovirus genome.";
RL   J. Gen. Virol. 84:17-28(2003).
DR   EMBL; AF480884; AAM00697.1; -; Genomic_DNA.
DR   InterPro; IPR006928; Herpes_teg_N.
DR   Pfam; PF04843; Herpes_teg_N; 1.
SQ   SEQUENCE   2275 AA;  255990 MW;  45BBA419CA576BCD CRC64;

  Query Match             100.0%;  Score 41;  DB 2;  Length 2275;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+03;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              |||||||
Db        277 RPLPIPP 283


RESULT 10
Q5VPS2_ORYSA
ID   Q5VPS2_ORYSA PRELIMINARY;      PRT;   175 AA.
AC   Q5VPS2;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Nucleoid DNA-binding protein cnd41-like.
GN   Name=OSJNBa0062J13.18;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC
RT   clone:OSJNBa0062J13.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP003564; BAD68553.1; -; Genomic_DNA.
DR   GO; GO:0003677; F:DNA binding; IEA.
KW   DNA-binding.
SQ   SEQUENCE   175 AA;  18553 MW;  C066FF696E2C786E CRC64;

  Query Match             97.6%;  Score 40;  DB 2;  Length 175;
  Best Local Similarity   85.7%;  Pred. No. 1.2e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||||:||
Db         43 RPLPVPP 49


RESULT 11
TLX2_HUMAN
ID   TLX2_HUMAN     STANDARD;      PRT;   284 AA.
AC   O43763; Q9UQ48;
DT   15-JUL-1999 (Rel. 38, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   T-cell leukemia homeobox protein 2 (Homeobox protein Hox-11L1) (Neural
DE   crest homeobox protein).
GN   Name=TLX2; Synonyms=HOX11L1, NCX;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Delgado P., Rodriguez R.E., Gonzalez-Sarmiento R.;
RT   "Genomic characterization and chromosomal location of the human
RT   homeobox gene HOX11L1.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99377083; PubMed=10446220; DOI=10.1074/jbc.274.34.24401;
RA   Iitsuka Y., Shimizu H., Kang M.M., Sasagawa K., Sekiya S.,
RA   Tokuhisa T., Hatano M.;
RT   "An enhancer element for expression of the Ncx (Enx, Hox11L1) gene in
RT   neural crest-derived cells.";
RL   J. Biol. Chem. 274:24401-24407(1999).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- SIMILARITY: Contains 1 homeobox DNA-binding domain.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AJ002607; CAA05636.1; -; Genomic_DNA.
DR   EMBL; AJ002608; CAA05636.1; JOINED; Genomic_DNA.
DR   EMBL; AJ002609; CAA05636.1; JOINED; Genomic_DNA.
DR   EMBL; AB008501; BAA83463.1; -; mRNA.
DR   EMBL; BC006356; AAH06356.1; -; mRNA.
DR   HSSP; P13297; 1IG7.
DR   TRANSFAC; T04367; -.
DR   Ensembl; ENSG00000115297; Homo sapiens.
DR   HGNC; HGNC:5057; TLX2.
DR   MIM; 604240; -.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR012287; Homeodomain-rel.
DR   Pfam; PF00046; Homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Developmental protein; DNA-binding; Homeobox; Nuclear protein.
FT   DNA_BIND    157    216       Homeobox.
FT   COMPBIAS     27    115       Gly-rich.
FT   CONFLICT     16     16       P -> A (in Ref. 1).
FT   CONFLICT     30     32       TPG -> PR (in Ref. 1).
FT   CONFLICT     37     48       LGRGGQGHGENG -> WVAGQVIGEWA (in Ref. 1).
FT   CONFLICT    100    102       Missing (in Ref. 1).
FT   CONFLICT    131    136       RLTAAL -> PAV (in Ref. 1).
FT   CONFLICT    219    219       Missing (in Ref. 1).
FT   CONFLICT    241    241       R -> T (in Ref. 1).
FT   CONFLICT    274    274       V -> A (in Ref. 1).
SQ   SEQUENCE   284 AA;  30251 MW;  794B07A9E7817939 CRC64;

  Query Match             97.6%;  Score 40;  DB 1;  Length 284;
  Best Local Similarity   85.7%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||||:||
Db         87 RPLPVPP 93


RESULT 12
TLX2_MOUSE
ID   TLX2_MOUSE     STANDARD;      PRT;   284 AA.
AC   Q61663;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   T-cell leukemia homeobox protein 2 (Homeobox protein Hox-11L1)
DE   (Homeobox TLX-2) (PMUR10F).
GN   Name=Tlx2; Synonyms=Hox1111, Tlx111;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=92020958; PubMed=1681546;
RA   Kennedy M.A., Gonzalez-Sarmiento R., Kees U.R., Lampert F., Dear T.N.,
RA   Boehm T., Rabbitts T.H.;
RT   "HOX11, a homeobox-containing T-cell oncogene on human chromosome
RT   10q24.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:8900-8904(1991).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- SIMILARITY: Contains 1 homeobox DNA-binding domain.
CC   -!- CAUTION: Was originally (Ref.1) thought to be the ortholog of
CC       human HOX11.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; M75953; AAA37805.1; -; mRNA.
DR   PIR; B41224; B41224.
DR   HSSP; P13297; 1IG7.
DR   TRANSFAC; T04368; -.
DR   Ensembl; ENSMUSG00000030040; Mus musculus.
DR   MGI; MGI:1350935; Tlx2.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR012287; Homeodomain-rel.
DR   Pfam; PF00046; Homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Developmental protein; DNA-binding; Homeobox; Nuclear protein.
FT   DNA_BIND    157    216       Homeobox.
FT   COMPBIAS     27    115       Gly-rich.
SQ   SEQUENCE   284 AA;  30361 MW;  CD1D6D3EB0F8CBDA CRC64;

  Query Match             97.6%;  Score 40;  DB 1;  Length 284;
  Best Local Similarity   85.7%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||||:||
```

```
Db         87 RPLPVPP 93


RESULT 13
Q8QRS5_9BETA
ID   Q8QRS5_9BETA PRELIMINARY;      PRT;   852 AA.
AC   Q8QRS5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Transcriptional transactivator TRS1.
OS   Pongine herpesvirus 4 (Chimpanzee cytomegalovirus).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
OX   NCBI_TaxID=188763;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;
RA   Davison A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
RA   Alcendor D.J., McGeoch D.J., Hayward G.S.;
RT   "The human cytomegalovirus genome revisited: comparison with the
RT   chimpanzee cytomegalovirus genome.";
RL   J. Gen. Virol. 84:17-28(2003).
DR   EMBL; AF480884; AAM00813.1; -; Genomic_DNA.
DR   InterPro; IPR003360; US22.
DR   Pfam; PF02393; US22; 1.
SQ   SEQUENCE   852 AA;  91991 MW;  9A9C27FDD94D1025 CRC64;

  Query Match             97.6%;  Score 40;  DB 2;  Length 852;
  Best Local Similarity   85.7%;  Pred. No. 6.8e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||||:||
Db        763 RPLPVPP 769


RESULT 14
Q4PG90_USTMA
ID   Q4PG90_USTMA PRELIMINARY;      PRT;  1074 AA.
AC   Q4PG90;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=UM00873.1;
OS   Ustilago maydis 521.
OC   Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC   Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX   NCBI_TaxID=237631;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=521;
RA   Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA   Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA   Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA   Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA   Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA   Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA   Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA   David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA   Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA   Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA   Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA   Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA   Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA   Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,
RA   Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA   Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA   Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA   Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA   Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA   Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA   Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA   Mozes J., Mulrain L., Munson G., Naylor J., Newes C., Nguyen C.,
RA   Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA   Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA   O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piqani B.,
RA   Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA   Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA   Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA   Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA   Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA   Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA   Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA   Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA   Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA   Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA   Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA   Zimmer A., Zody M., Lander E.;
RT   "The genome sequence of Ustilago maydis.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACP01000028; EAK81623.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1074 AA;  106245 MW;  25B95F6F96C91F80 CRC64;

  Query Match             97.6%;  Score 40;  DB 2;  Length 1074;
  Best Local Similarity   85.7%;  Pred. No. 8.7e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||||:||
Db        661 RPLPVPP 667


RESULT 15
Q993Z6_HPV26
ID   Q993Z6_HPV26 PRELIMINARY;      PRT;    88 AA.
AC   Q993Z6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative E4 protein.
OS   Human papillomavirus - 82.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Alphapapillomavirus.
OX   NCBI_TaxID=129724;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Terai M., Burk R.D.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF293961; AAK28453.1; -; Genomic_DNA.
DR   InterPro; IPR003861; Papilloma_E4.
DR   Pfam; PF02711; Pap_E4; 1.
SQ   SEQUENCE   88 AA;  10084 MW;  6752D8CF3A9475D7 CRC64;

  Query Match             95.1%;  Score 39;  DB 2;  Length 88;
  Best Local Similarity   85.7%;  Pred. No. 85;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPIPP 7
              ||:||||
Db         24 RPIPIPP 30


Search completed: April  4, 2006, 13:15:18
Job time : 8.35079 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 3.47251 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-290
Perfect score: 39
Sequence: 1 RPLPSRP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	100.0	7	3	AAB17234	Aab17234 SH3 antag
2	39	100.0	7	5	ABB73227	Abb73227 Src homol
3	39	100.0	7	7	ADJ73381	Adj73381 SH3 antag
4	39	100.0	7	8	ADJ53015	Adj53015 CH1 delet
5	39	100.0	7	8	ADJ51976	Adj51976 CH1 delet
6	39	100.0	13	2	AAW11123	Aaw11123 Src SH3 d
7	39	100.0	50	2	AAW16936	Aaw16936 Random re
8	39	100.0	50	2	AAW25499	Aaw25499 Random pe
9	39	100.0	79	4	AAU48480	Aau48480 Propionib
10	39	100.0	79	6	ABM44999	Abm44999 Propionib
11	39	100.0	85	4	AAU50101	Aau50101 Propionib
12	39	100.0	85	6	ABM46620	Abm46620 Propionib
13	39	100.0	199	4	ABG07608	Abg07608 Novel hum
14	39	100.0	351	2	AAW72022	Aaw72022 HSV-2 str
15	36	92.3	7	3	AAB17235	Aab17235 SH3 antag
16	36	92.3	7	5	ABB73228	Abb73228 Src homol
17	36	92.3	7	7	ADJ73382	Adj73382 SH3 antag
18	36	92.3	7	8	ADJ53016	Adj53016 CH1 delet
19	36	92.3	7	8	ADJ51977	Adj51977 CH1 delet
20	36	92.3	13	2	AAW11116	Aaw11116 Src SH3 d
21	36	92.3	34	2	AAW25498	Aaw25498 Random pe
22	36	92.3	35	2	AAW16935	Aaw16935 Random re
23	36	92.3	70	7	ABO77807	Abo77807 Pseudomon
24	36	92.3	76	8	ADR94375	Adr94375 Novel S.
25	36	92.3	76	9	AEA58245	Aea58245 Streptoco
26	36	92.3	116	8	ADQ67167	Adq67167 Novel hum
27	36	92.3	165	4	AAB82136	Aab82136 Human sbg
28	36	92.3	257	7	ABO84172	Abo84172 Pseudomon
29	36	92.3	280	4	ABG05901	Abg05901 Novel hum
30	36	92.3	455	7	ABM89085	Abm89085 Rice abio
31	36	92.3	476	4	ABG28212	Abg28212 Novel hum
32	36	92.3	507	8	ADS21668	Ads21668 Bacterial
33	36	92.3	710	9	ABM92205	Abm92205 M. xanthu
34	36	92.3	722	7	ADD46839	Add46839 Rat Prote
35	36	92.3	722	7	ADE56292	Ade56292 Rat Prote
36	36	92.3	728	4	AAM78754	Aam78754 Human pro
37	36	92.3	728	7	ADD46841	Add46841 Human Pro
38	36	92.3	728	7	ADE56294	Ade56294 Human Pro
39	36	92.3	728	8	ADJ66562	Adj66562 PI3 kinas
40	36	92.3	729	7	ADJ71151	Adj71151 Human hea
41	36	92.3	743	4	AAM79738	Aam79738 Human pro
42	36	92.3	1138	3	AAY83222	Aay83222 CAP6 poly
43	36	92.3	1138	7	ADJ93590	Adj93590 Mouse bon
44	35	89.7	61	4	AAU61321	Aau61321 Propionib
45	35	89.7	61	6	ABM57840	Abm57840 Propionib

ALIGNMENTS

RESULT 1
AAB17234
ID AAB17234 standard; peptide; 7 AA.
XX
AC AAB17234;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:290.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 298; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 39; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 1 RPLPSRP 7

RESULT 2
ABB73227
ID ABB73227 standard; peptide; 7 AA.
XX
AC ABB73227;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:290.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 39; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 1 RPLPSRP 7

RESULT 3
ADJ73381
ID ADJ73381 standard; peptide; 7 AA.
XX
AC ADJ73381;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 836.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 836; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 39; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 1 RPLPSRP 7

RESULT 4
ADJ53015
ID ADJ53015 standard; peptide; 7 AA.
XX
AC ADJ53015;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID836.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 836; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 39; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 1 RPLPSRP 7

RESULT 5
ADJ51976
ID ADJ51976 standard; peptide; 7 AA.
XX
AC ADJ51976;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID836.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 836; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be

```
CC   useful for the development of compounds with an osteopathic,
CC   cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC   gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC   immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC   antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC   respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC   modulator or cytokine-agonist. The methods and compositions of the
CC   present invention are useful for the diagnosis, prevention and/or
CC   treatment of diseases or conditions associated with aberrant expression
CC   or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC   cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC   endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC   obstetric, haematologic, immunological, allergic, infectious,
CC   musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC   pediatric, psychiatric, renal or pulmonary disorders. The present
CC   sequence is that of a peptide which may be used during the creation of a
CC   mimetibody of the invention.
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 39;  DB 8;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              |||||||
Db          1 RPLPSRP 7


RESULT 6
AAW11123
ID   AAW11123 standard; peptide; 13 AA.
XX
AC   AAW11123;
XX
DT   25-JUN-1997  (first entry)
XX
DE   Src SH3 domain-binding peptide used in signal transduction modulation.
XX
KW   Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW   protein tyrosine kinase; signal transduction; RNA processing;
KW   trafficking; translation.
XX
OS   Synthetic.
XX
PN   WO9603649-A1.
XX
PD   08-FEB-1996.
XX
PF   24-JUL-1995;   95WO-US009382.
XX
PR   22-JUL-1994;   94US-00278865.
PR   07-JUN-1995;   95US-00483555.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ;
XX
DR   WPI; 1996-117151/12.
XX
PT   Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT   proteins - useful for e.g. modulating signal transduction pathways at the
PT   cellular level, esp. protein tyrosine kinase-mediated.
XX
PS   Claim 40; Page 84; 116pp; English.
XX
CC   AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC   binding peptides are useful in modulating signal transduction pathways at
CC   the cellular level (especially protein tyrosine kinase-mediated),
CC   modulating oncogenic protein activity, or providing compounds for the
CC   development of drugs with the ability to modulate broad classes, as well
CC   as specific classes, of proteins involved in signal transduction and also
CC   for regulating the processing, trafficking or translation of RNA.
CC   Conjugates of the peptides with detectable labels or imaging agents are
CC   useful for imaging cells, tissues and organs in which Src or Src-related
CC   proteins are expressed
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 39;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              |||||||
Db          4 RPLPSRP 10


RESULT 7
AAW16936
ID   AAW16936 standard; peptide; 50 AA.
XX
AC   AAW16936;
XX
DT   27-JUN-1997  (first entry)
XX
DE   Random recombinant SH3 domain binding peptide.
XX
KW   Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW   protein tyrosine kinase; signal transduction; RNA processing;
KW   trafficking; translation.
XX
OS   Synthetic.
XX
PN   WO9603649-A1.
XX
PD   08-FEB-1996.
XX
PF   24-JUL-1995;   95WO-US009382.
XX
PR   22-JUL-1994;   94US-00278865.
PR   07-JUN-1995;   95US-00483555.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ;
XX
DR   WPI; 1996-117151/12.
XX
PT   Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT   proteins - useful for e.g. modulating signal transduction pathways at the
PT   cellular level, esp. protein tyrosine kinase-mediated.
XX
PS   Disclosure; Fig 1; 116pp; English.
XX
CC   AAW16924-W16948 are random recombinant peptides derived from one of three
CC   peptide libraries, T9, T12 and R8C. The peptides are all SH3 domain-
CC   binding peptides. SH3 binding peptides are useful in modulating signal
CC   transduction pathways at the cellular level (especially protein tyrosine
CC   kinase-mediated), modulating oncogenic protein activity, or providing
CC   compounds for the development of drugs with the ability to modulate broad
CC   classes, as well as specific classes, of proteins involved in signal
CC   transduction and also for regulating the processing, trafficking or
CC   translation of RNA. Conjugates of the peptides with detectable labels or
CC   imaging agents are useful for imaging cells, tissues and organs in which
CC   Src or Src-related proteins are expressed
XX
SQ   Sequence 50 AA;

  Query Match             100.0%;  Score 39;  DB 2;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              |||||||
Db         37 RPLPSRP 43
```

RESULT 8
AAW25499
ID AAW25499 standard; peptide; 50 AA.
XX
AC AAW25499;
XX
DT 27-MAR-1998 (first entry)
XX
DE Random peptide recombinant clone T9.SRC3.3.
XX
KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS Synthetic.
OS Unidentified.
XX
PN WO9730074-A1.
XX
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-US002298.
XX
PR 16-FEB-1996; 96US-00602999.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;
XX
DR WPI; 1997-424972/39.
XX
PT Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.
XX
PS Disclosure; Fig 5; 131pp; English.
XX
CC The present sequence represents a random peptide recombinant isolated by
CC the method of the present invention. SH3 (Src homology region 3) binding
CC peptides are selected from: (a) peptides which bind the SH3 domain of
CC Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)
CC peptides which bind the SH3 domain of Abl; (d) peptides which bind the
CC SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;
CC (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind
CC the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3
CC domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain
CC of Grb2. The purified binding peptides can be used in the method to
CC identify inhibitors of their binding to their respective SH3 domains,
CC which could be used to modulate the pharmacological activity of proteins
CC or polypeptide containing the SH3 domain. The peptides can also be used
CC to activate Src or Src-related protein tyrosine kinases, to stimulate the
CC immune response by increasing the production of certain lymphokines, e.g.
CC tumour necrosis factor-alpha and interleukin-1, or to deliver a
CC conjugated molecule to certain cellular compartments containing Src or
CC Src related proteins
XX
SQ Sequence 50 AA;

Query Match 100.0%; Score 39; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 37 RPLPSRP 43

RESULT 9
AAU48480
ID AAU48480 standard; protein; 79 AA.
XX
AC AAU48480;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #9376.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59542.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 9675; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 79 AA;

Query Match 100.0%; Score 39; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 54 RPLPSRP 60

RESULT 10
ABM44999

ID ABM44999 standard; protein; 79 AA.
XX
AC ABM44999;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #9675.
XX
KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64471.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 9675; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 79 AA;

Query Match 100.0%; Score 39; DB 6; Length 79;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 54 RPLPSRP 60

RESULT 11
AAU50101
ID AAU50101 standard; protein; 85 AA.
XX
AC AAU50101;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #10997.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59546.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 11296; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 85 AA;

Query Match 100.0%; Score 39; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 27 RPLPSRP 33

RESULT 12
ABM46620
ID ABM46620 standard; protein; 85 AA.
XX
AC ABM46620;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #11296.
XX
KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
DR N-PSDB; ACF64475.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Example 1; SEQ ID NO 11296; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 85 AA;

Query Match 100.0%; Score 39; DB 6; Length 85;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7
|||||||
Db 27 RPLPSRP 33

RESULT 13
ABG07608
ID ABG07608 standard; protein; 199 AA.
XX
AC ABG07608;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #7599.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS71795.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 37967; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 199 AA;

Query Match 100.0%; Score 39; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPLPSRP 7

```
Db          1 RPLPSRP 7


RESULT 14
AAW72022
ID   AAW72022 standard; protein; 351 AA.
XX
AC   AAW72022;
XX
DT   07-DEC-1998  (first entry)
XX
DE   HSV-2 strain SB5 Contig ID 102 ORF#8 protein.
XX
KW   HSV-2 strain SB5; immunological response induction; therapy;
KW   antiviral identification; viral protein inhibitor.
XX
OS   Herpes simplex virus 2.
XX
PN   WO9820016-A1.
XX
PD   14-MAY-1998.
XX
PF   31-OCT-1997;   97WO-US020016.
XX
PR   04-NOV-1996;   96US-0030279P.
PR   09-JUN-1997;   97US-0049018P.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Esser KM,  Chan JY,  Dabrowski-Amaral CE,  Delvecchio AM,  Dillon SB;
PI   Leary JJ;
XX
DR   WPI; 1998-286847/25.
DR   N-PSDB; AAV62132.
XX
PT   Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
PT   treatment of infection or inducing immunological response in mammal.
XX
PS   Claim 10; Page 47; 748pp; English.
XX
CC   This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC   sequence of the invention. This sequence was isolated from a HSV-2 strain
CC   SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 102.
CC   The proteins can be used for the treatment or prevention of disease, to
CC   induce an immunological response in a mammal or to identify inhibitors,
CC   activators or novel antivirals. Antagonists of the proteins can be used
CC   to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC   it can also be used to induce an immunological response in a mammal
XX
SQ   Sequence 351 AA;

  Query Match             100.0%;  Score 39;  DB 2;  Length 351;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              |||||||
Db        284 RPLPSRP 290


RESULT 15
AAB17235
ID   AAB17235 standard; peptide; 7 AA.
XX
AC   AAB17235;
XX
DT   31-OCT-2000  (first entry)
XX
DE   SH3 antagonist peptide sequence SEQ ID NO:291.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW   inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW   thrombosis; pharmaceutical.
XX
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-US025044.
XX
PR   23-OCT-1998;   98US-0105371P.
PR   22-OCT-1999;   99US-00428082.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham J,  Boone TC;
XX
DR   WPI; 2000-350702/30.
XX
PT   Novel composition of matter comprising an Fc domain and pharmacologically
PT   active peptides, useful for treating cancer and autoimmune diseases.
XX
PS   Claim 39; Page 298; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC   (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC   P3, and P4 = are each independently sequences of pharmacologically active
CC   peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC   c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC   of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC   thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC   cells from the present invention can be used for producing pharmaceutical
CC   compositions. The compositions are useful for treating cancer, asthma,
CC   thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC   a Fab domain) can provide a longer half-life or incorporate functions
CC   such as Fc receptor binding, protein A binding, complement fixation, and
CC   possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC   AAB18003 represent nucleotide and amino acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             92.3%;  Score 36;  DB 3;  Length 7;
  Best Local Similarity   85.7%;  Pred. No. 2e+06;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db          1 RPLPTRP 7


Search completed: April  4, 2006, 13:07:46
Job time : 4.47251 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 1.14529 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-290
Perfect score: 39
Sequence: 1 RPLPSRP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	100.0	1751	2	T09394	gag-pro-pol polypr
2	36	92.3	103	2	B87261	hypothetical prote
3	36	92.3	270	2	C87568	transcription regu
4	36	92.3	369	2	AG1950	hypothetical prote
5	36	92.3	723	2	B38749	3-phosphatidylinos
6	36	92.3	728	2	H59435	phosphoinositide-3
7	35	89.7	175	2	AH2827	dihydrofolate redu
8	35	89.7	175	2	F97605	dihydrofolate redu
9	35	89.7	208	2	S27657	hypothetical prote
10	35	89.7	240	2	A83462	hypothetical prote
11	35	89.7	828	2	T03544	hypothetical prote
12	34	87.2	127	2	F72561	hypothetical prote
13	34	87.2	129	2	T21290	hypothetical prote
14	34	87.2	164	2	F72470	hypothetical prote
15	34	87.2	179	2	F83305	hypothetical prote
16	34	87.2	187	2	T27416	hypothetical prote
17	34	87.2	209	2	T27030	hypothetical prote
18	34	87.2	235	2	A72594	hypothetical prote
19	34	87.2	309	2	G87498	hypothetical prote
20	34	87.2	312	2	A61183	hypothetical prote
21	34	87.2	326	2	AE3125	mannonate dehydrat
22	34	87.2	326	2	D98162	mannonate dehydrat
23	34	87.2	421	1	S11674	acrosin (EC 3.4.21
24	34	87.2	462	1	QQBED4	HHRF4 protein - hu
25	34	87.2	588	2	T24980	hypothetical prote
26	34	87.2	661	2	T22319	hypothetical prote
27	34	87.2	953	2	T40643	probable serine th
28	34	87.2	1119	2	T50995	related to cytoske
29	34	87.2	1238	1	JC5573	copper-transportin
30	34	87.2	1440	2	T27942	lin-15B protein -
31	34	87.2	1611	2	T38236	hypothetical prote
32	33	84.6	75	2	S05589	Balbiani ring prot
33	33	84.6	82	2	C48349	UL28 protein - sai
34	33	84.6	85	2	S10120	Balbiani ring prot
35	33	84.6	85	2	S10119	Balbiani ring prot
36	33	84.6	196	2	I76912	ychG protein - Esc
37	33	84.6	227	2	B83505	hypothetical prote
38	33	84.6	233	2	D95877	probable transcrip
39	33	84.6	278	1	TPHUTW	troponin T, slow s
40	33	84.6	314	2	T03775	DNA-binding homeot
41	33	84.6	317	2	E86264	protein F3F19.7 [i
42	33	84.6	348	2	T33179	hypothetical prote
43	33	84.6	432	2	A25483	env polyprotein, r
44	33	84.6	451	2	B81850	exonuclease VII la
45	33	84.6	459	2	S33000	hypothetical prote

ALIGNMENTS

RESULT 1
T09394
gag-pro-pol polyprotein - walleye dermal sarcoma virus
C;Species: walleye dermal sarcoma virus
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C;Accession: T09394; T09393
R;Petropoulos, C.J.
submitted to the EMBL Data Library, November 1997
A;Description: Appendix 2: Retroviral taxonomy, protein structure, sequences, and gene
A;Reference number: Z16660
A;Accession: T09394
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
A;Residues: 1-1751 <PET>
A;Cross-references: UNIPROT:O92815; UNIPARC:UPI000010BDA0; EMBL:AF033822; NID:g2801519
A;Accession: T09393
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
A;Residues: 1-582 <PEW>
A;Cross-references: UNIPARC:UPI000010499D; EMBL:AF033822; NID:g2801519; PID:g2801521
C;Genetics:
A;Gene: gag-pro-pol
A;Introns: 582/3

Query Match 100.0%; Score 39; DB 2; Length 1751;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              |||||||
Db       1479 RPLPSRP 1485
```

RESULT 2
B87261
hypothetical protein CC0099 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87261
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87261
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <STO>
A;Cross-references: UNIPROT:Q9ABX1; UNIPARC:UPI00000C6F22; GB:AE005673; NID:g13421202;
C;Genetics:
A;Gene: CC0099

```
  Query Match             92.3%;  Score 36;  DB 2;  Length 103;
  Best Local Similarity   85.7%;  Pred. No. 15;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db         10 RPLPNRP 16


RESULT 3
C87568
transcription regulator, AraC family [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: C87568
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87568
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <STO>
A;Cross-references: UNIPROT:Q9A584; UNIPARC:UPI00000C7789; GB:AE005673; NID:g13424141; P
C;Genetics:
A;Gene: CC2573

  Query Match             92.3%;  Score 36;  DB 2;  Length 270;
  Best Local Similarity   85.7%;  Pred. No. 40;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db         34 RPLPARP 40


RESULT 4
AG1950
hypothetical protein all1154 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AG1950
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG1950
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-369 <KUR>
A;Cross-references: UNIPROT:Q8YXQ8; UNIPARC:UPI00000CDFD1; GB:BA000019; PIDN:BAB73111.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1154

  Query Match             92.3%;  Score 36;  DB 2;  Length 369;
  Best Local Similarity   85.7%;  Pred. No. 54;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db        168 RPLPTRP 174


RESULT 5
B38749
3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain B - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 15-Mar-2004
C;Accession: B38749
R;Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, A

Cell 65, 91-104, 1991
A;Title: Characterization of two 85 kd proteins that associate with receptor tyrosine ki
A;Reference number: A38749; MUID:91191567; PMID:1707345
A;Accession: B38749
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-723 <OTS>
A;Cross-references: UNIPARC:UPI000017C48F; GB:M61745; GB:M61746
C;Keywords: phosphotransferase
F;325-420/Domain: SH2 homology <SH2A>
F;617-706/Domain: SH2 homology <SH2>

  Query Match             92.3%;  Score 36;  DB 2;  Length 723;
  Best Local Similarity   85.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db         94 RPLPARP 100


RESULT 6
H59435
phosphoinositide-3-kinase regulatory beta chain [imported] - human
C;Species: Homo sapiens (man)
C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C;Accession: H59435; A59436
R;Volinia S; Patracchini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L;
Oncogene 7, 789-793, 1992
A;Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinositol
A;Reference number: H59435
A;Accession: H59435
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-728 <VOL>
A;Cross-references: UNIPROT:O00459; UNIPARC:UPI000013106C; GB:NP_005018; PID:g4826908;
R;Janssen, J.W.; Schleithoff, L.; Bartram, C.R.; Schulz, A.S.
Oncogene 16, 1767-1772, 1998
A;Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subun
A;Reference number: A59436; MUID:98241181; PMID:9582025
A;Accession: A59436
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-728 <JAN>
A;Cross-references: UNIPARC:UPI000013106C; GB:NP_005018; PID:g4826908; PIDN:NP_005018.1

  Query Match             92.3%;  Score 36;  DB 2;  Length 728;
  Best Local Similarity   85.7%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db         94 RPLPARP 100


RESULT 7
AH2827
dihydrofolate reductase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 05-Oct-2004
C;Accession: AH2827
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClel
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
```

A;Accession: AH2827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-175 <KUR>
A;Cross-references: UNIPROT:Q8UDS4; UNIPARC:UPI00000D1D73; GB:AE008688; PIDN:AAL43038.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: folA
A;Map position: circular chromosome
C;Superfamily: dihydrofolate reductase; type I dihydrofolate reductase homology

Query Match 89.7%; Score 35; DB 2; Length 175;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              |||| ||
Db         57 RPLPGRP 63
```

RESULT 8
F97605
dihydrofolate reductase (AP001518) [imported] - Agrobacterium tumefaciens (strain C58, C
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 05-Oct-2004
C;Accession: F97605
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F97605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-175 <KUR>
A;Cross-references: UNIPROT:Q8UDS4; UNIPARC:UPI00000D1D73; GB:AE007869; PIDN:AAK87799.1;
C;Genetics:
A;Gene: AGR_C_3708
A;Map position: circular chromosome
C;Superfamily: dihydrofolate reductase; type I dihydrofolate reductase homology

Query Match 89.7%; Score 35; DB 2; Length 175;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              |||| ||
Db         57 RPLPGRP 63
```

RESULT 9
S27657
hypothetical protein 1 - Rhizobium meliloti
C;Species: Rhizobium meliloti
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S27657
R;Miller, K.J.; McKinstry, M.W.; Hunt, W.P.; Nixon, B.
submitted to the EMBL Data Library, May 1992
A;Description: Identification of the diglyceride kinase structural gene of Rhizobium mel
A;Reference number: S27657
A;Accession: S27657
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-208 <MIL>
A;Cross-references: UNIPROT:Q52921; UNIPARC:UPI00000B5865; EMBL:M94085; NID:g152176; PID

Query Match 89.7%; Score 35; DB 2; Length 208;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              |||| ||
Db         86 RPLPDRP 92
```

RESULT 10
A83462
hypothetical protein PA1469 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A83462
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Li
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83462
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-240 <STO>
A;Cross-references: UNIPROT:Q9I3P3; UNIPARC:UPI00000C5390; GB:AE004576; GB:AE004091; N
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1469
C;Superfamily: Streptomyces coelicolor hypothetical protein SC4A10.14c

Query Match 89.7%; Score 35; DB 2; Length 240;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              |||| ||
Db        159 RPLPERP 165
```

RESULT 11
T03544
hypothetical protein - Rhodobacter capsulatus
C;Species: Rhodobacter capsulatus
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03544
R;Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.
Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997
A;Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB10
A;Reference number: Z14955; MUID:97404404; PMID:9256491
A;Accession: T03544
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-828 <VLC>
A;Cross-references: UNIPROT:O68107; UNIPARC:UPI00000BCF2D; EMBL:AF010496; NID:g3128256
C;Genetics:
A;Map position: 1

Query Match 89.7%; Score 35; DB 2; Length 828;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              :||||||
Db        453 QPLPSRP 459
```

RESULT 12
F72561
hypothetical protein APE1776 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F72561
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tak
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72561
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-127 <KAW>
A;Cross-references: UNIPROT:Q9YB19; UNIPARC:UPI000005E0A7; DDBJ:AP000062; NID:g5105244;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1776

Query Match 87.2%; Score 34; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 PLPSRP 7
              ||||||
Db         30 PLPSRP 35
```

RESULT 13
T21290
hypothetical protein F23B12.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T21290
R;Wild, A.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z19402
A;Accession: T21290
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-129 <WIL>
A;Cross-references: UNIPROT:Q19751; UNIPARC:UPI000007FB84; EMBL:Z77659; PIDN:CAB01166.1;
A;Experimental source: clone F23B12
C;Genetics:
A;Gene: CESP:F23B12.4
A;Map position: 5
A;Introns: 26/2; 38/3; 89/1

Query Match 87.2%; Score 34; DB 2; Length 129;
Best Local Similarity 71.4%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              ||:|:||
Db         34 RPIPARP 40
```

RESULT 14
F72470
hypothetical protein APE2407 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F72470
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72470
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-164 <KAW>
A;Cross-references: UNIPROT:Q9Y978; UNIPARC:UPI000005E32A; DDBJ:AP000064; NID:g5105945;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2407

Query Match 87.2%; Score 34; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 PLPSRP 7
              ||||||
Db         89 PLPSRP 94
```

RESULT 15
F83305
hypothetical protein PA2724 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83305
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83305
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-179 <STO>
A;Cross-references: UNIPROT:Q9I0B8; UNIPARC:UPI00000C57D0; GB:AE004700; GB:AE004091; NI
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2724

Query Match 87.2%; Score 34; DB 2; Length 179;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              |||| ||
Db        142 RPLPPRP 148
```

Search completed: April 4, 2006, 13:17:29
Job time : 2.14529 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 5.35079 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-290
Perfect score: 39
Sequence: 1 RPLPSRP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	100.0	313	2	Q4NSC7_9DELT	Q4nsc7 anaeromyxob
2	39	100.0	847	2	Q59FE5_HUMAN	Q59fe5 homo sapien
3	39	100.0	1126	2	Q5YNA3_NOCFA	Q5yna3 nocardia fa
4	39	100.0	1751	2	O92815_9RETR	O92815 walleye der
5	37	94.9	169	2	Q5GXW8_XANOR	Q5gxw8 xanthomonas
6	37	94.9	320	2	Q35392_9AVES	Q35392 phaenicopha
7	36	92.3	103	2	Q9ABX1_CAUCR	Q9abx1 caulobacter
8	36	92.3	168	2	Q4HB55_9DEIO	Q4hb55 deinococcus
9	36	92.3	208	2	Q5SIN6_THET8	Q5sin6 thermus the
10	36	92.3	270	2	Q9A584_CAUCR	Q9a584 caulobacter
11	36	92.3	313	2	Q5NTG0_9BACT	Q5ntg0 uncultured
12	36	92.3	334	2	Q8PLV6_XANAC	Q8plv6 xanthomonas
13	36	92.3	352	2	Q84SM1_ORYSA	Q84sm1 oryza sativ
14	36	92.3	369	2	Q8YXQ8_ANASP	Q8yxq8 anabaena sp
15	36	92.3	479	2	Q72GL1_THET2	Q72gl1 thermus the
16	36	92.3	485	2	Q67KW2_SYMTH	Q67kw2 symbiobacte
17	36	92.3	495	2	Q5KBL0_CRYNE	Q5kbl0 cryptococcu
18	36	92.3	501	2	Q67M24_SYMTH	Q67m24 symbiobacte
19	36	92.3	504	2	Q67ST3_SYMTH	Q67st3 symbiobacte
20	36	92.3	625	2	Q4G045_RAT	Q4g045 rattus norv
21	36	92.3	661	2	Q55N79_CRYNE	Q55n79 cryptococcu
22	36	92.3	722	1	P85B_MOUSE	O08908 mus musculu
23	36	92.3	722	1	P85B_RAT	Q63788 rattus norv
24	36	92.3	722	2	Q5FVS6_RAT	Q5fvs6 rattus norv
25	36	92.3	722	2	Q5U3K7_MOUSE	Q5u3k7 mus musculu
26	36	92.3	724	1	P85B_BOVIN	P23726 bos taurus
27	36	92.3	724	2	Q8XT28_RALSO	Q8xt28 ralstonia s
28	36	92.3	728	1	P85B_HUMAN	O00459 homo sapien
29	36	92.3	728	2	Q5EAT5_HUMAN	Q5eat5 homo sapien
30	36	92.3	776	2	Q5MB23_BRARE	Q5mb23 brachydanio
31	36	92.3	776	2	Q68EH4_BRARE	Q68eh4 brachydanio
32	36	92.3	776	2	Q6RI23_BRARE	Q6ri23 brachydanio
33	36	92.3	785	2	Q75UA3_FUGRU	Q75ua3 fugu rubrip
34	36	92.3	796	2	Q60DA0_ORYSA	Q60da0 oryza sativ
35	36	92.3	1014	2	Q4NUP3_9DELT	Q4nup3 anaeromyxob
36	36	92.3	1138	1	BMP2K_MOUSE	Q91z96 mus musculu
37	35	89.7	141	2	Q5SHZ0_THET8	Q5shz0 thermus the
38	35	89.7	141	2	Q72IA6_THET2	Q72ia6 thermus the
39	35	89.7	169	2	Q5DF29_SCHJA	Q5df29 schistosoma
40	35	89.7	175	2	Q8UDS4_AGRT5	Q8uds4 agrobacteri
41	35	89.7	180	2	Q92NQ7_RHIME	Q92nq7 rhizobium m
42	35	89.7	186	2	Q7XY65_GRIJA	Q7xy65 griffithsia
43	35	89.7	188	2	Q72JV9_THET2	Q72jv9 thermus the
44	35	89.7	208	2	Q52921_RHIME	Q52921 rhizobium m
45	35	89.7	235	2	Q4H1D0_9ACTO	Q4h1d0 actinoplane

ALIGNMENTS

RESULT 1
Q4NSC7_9DELT
ID Q4NSC7_9DELT PRELIMINARY; PRT; 313 AA.
AC Q4NSC7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE GAF.
GN ORFNames=AdehDRAFT_1899;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000022; EAL78543.1; -; Genomic_DNA.
SQ SEQUENCE 313 AA; 32405 MW; 5DC496DA9F40DDEE CRC64;

Query Match 100.0%; Score 39; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RPLPSRP 7
              |||||||
Db        129 RPLPSRP 135
```

RESULT 2
Q59FE5_HUMAN
ID Q59FE5_HUMAN PRELIMINARY; PRT; 847 AA.
AC Q59FE5;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE A disintegrin-like and metalloprotease (Reprolysin type) with
DE thrombospondin type 1 motif, 10 preproprotein variant (Fragment).

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RA   Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA   Ohara O., Nagase T., Kikuno F.R.;
RT   "None Title.";
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB209515; BAD92752.1; -; mRNA.
KW   Integrin; Metalloprotease; Protease.
FT   NON_TER       1      1
FT   NON_TER     847    847
SQ   SEQUENCE   847 AA;  91651 MW;  5E64B143620CE84F CRC64;

  Query Match             100.0%;  Score 39;  DB 2;  Length 847;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              |||||||
Db        712 RPLPSRP 718


RESULT 3
Q5YNA3_NOCFA
ID   Q5YNA3_NOCFA PRELIMINARY;      PRT;  1126 AA.
AC   Q5YNA3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Putative membrane protein.
GN   OrderedLocusNames=nfa54860;
OS   Nocardia farcinica.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Nocardiaceae; Nocardia.
OX   NCBI_TaxID=37329;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=IFM 10152;
RX   PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA   Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA   Shiba T., Hattori M.;
RT   "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR   EMBL; AP006618; BAD60338.1; -; Genomic_DNA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR004869; MMPL.
DR   Pfam; PF03176; MMPL; 1.
KW   Complete proteome.
SQ   SEQUENCE   1126 AA;  118284 MW;  0651A04BDC7FB9D0 CRC64;

  Query Match             100.0%;  Score 39;  DB 2;  Length 1126;
  Best Local Similarity   100.0%;  Pred. No. 4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              |||||||
Db        868 RPLPSRP 874


RESULT 4
O92815_9RETR
ID   O92815_9RETR PRELIMINARY;      PRT;  1751 AA.
AC   O92815;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Pr gag-pro-pol.
GN   Name=gag-pro-pol;
OS   Walleye dermal sarcoma virus.
OC   Viruses; Retroid viruses; Retroviridae; Epsilonretrovirus.
OX   NCBI_TaxID=39720;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Petropoulos C.J.;
RT   "Appendix 2: Retroviral taxonomy, protein structure, sequences, and
RT   genetic maps.";
RL   (In) Coffin J.M. (eds.);
RL   RETROVIRUSES, pp.757-0, Cold Spring Harbor Laboratory Press, Cold
RL   Spring Harbor, New York, NY, USA (1997).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Chappey C.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF033822; AAC82611.1; -; Genomic_RNA.
DR   PIR; T09394; T09394.
DR   HSSP; P03355; 1I6J.
DR   GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0004523; F:ribonuclease H activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.
DR   GO; GO:0006310; P:DNA recombination; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   GO; GO:0006278; P:RNA-dependent DNA replication; IEA.
DR   InterPro; IPR002156; RNaseH.
DR   InterPro; IPR001584; Rve.
DR   InterPro; IPR000477; RVTse.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00075; RnaseH; 1.
DR   Pfam; PF00665; rve; 1.
DR   Pfam; PF00078; RVT_1; 1.
DR   Pfam; PF00098; zf-CCHC; 1.
DR   SMART; SM00343; ZnF_C2HC; 1.
DR   PROSITE; PS50879; RNASE_H; 1.
DR   PROSITE; PS50158; ZF_CCHC; 1.
SQ   SEQUENCE   1751 AA;  196152 MW;  DB9561C775A12217 CRC64;

  Query Match             100.0%;  Score 39;  DB 2;  Length 1751;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              |||||||
Db       1479 RPLPSRP 1485


RESULT 5
Q5GXW8_XANOR
ID   Q5GXW8_XANOR PRELIMINARY;      PRT;   169 AA.
AC   Q5GXW8;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   PilX.
GN   Name=pilX; OrderedLocusNames=XOO3199;
OS   Xanthomonas oryzae (pv. oryzae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=64187;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=KACC10331 / KXO85;
RX   PubMed=15673718; DOI=10.1093/nar/gki206;
RA   Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA   Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
RA   Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA   Go S.-J.;
RT   "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
RT   the bacterial blight pathogen of rice.";
RL   Nucleic Acids Res. 33:577-586(2005).
DR   EMBL; AE013598; AAW76453.1; -; Genomic_DNA.
```

```
KW   Complete proteome.
SQ   SEQUENCE   169 AA;  18026 MW;  CA263FFDCB404F68 CRC64;

  Query Match             94.9%;  Score 37;  DB 2;  Length 169;
  Best Local Similarity   85.7%;  Pred. No. 1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||:||||
Db          8 RPMPSRP 14


RESULT 6
Q35392_9AVES
ID   Q35392_9AVES PRELIMINARY;      PRT;   320 AA.
AC   Q35392;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Cytochrome b (Fragment).
OS   Phaenicophaeus curvirostris.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Cuculiformes; Cuculidae;
OC   Phaenicophaeus.
OX   NCBI_TaxID=33595;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=94356264; PubMed=8075835; DOI=10.1006/mpev.1994.1019;
RA   Avise J.C., Nelson W.S., Sibley C.G.;
RT   "Why one-kilobase sequences from mitochondrial DNA fail to solve the
RT   Hoatzin phylogenetic enigma.";
RL   Mol. Phylogenet. Evol. 3:175-184(1994).
CC   -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC       complex (complex III or cytochrome b-c1 complex), which is a
CC       respiratory chain that generates an electrochemical potential
CC       coupled to ATP synthesis (By similarity).
CC   -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC   -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC       cytochrome c1 and the Rieske protein (By similarity).
CC   -!- SIMILARITY: Belongs to the cytochrome b family.
DR   EMBL; U09264; AAA65036.1; -; Genomic_DNA.
DR   SMR; Q35392; 1-320.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR   GO; GO:0005739; C:mitochondrion; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR005798; Cytb_b6_C.
DR   InterPro; IPR005797; Cytb_b6_N.
DR   Pfam; PF00032; Cytochrom_B_C; 1.
DR   Pfam; PF00033; Cytochrom_B_N; 1.
DR   PROSITE; PS51003; CYTB_CTER; 1.
DR   PROSITE; PS51002; CYTB_NTER; 1.
KW   Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW   Respiratory chain; Transmembrane; Transport.
FT   NON_TER       1      1
FT   NON_TER     320    320
SQ   SEQUENCE   320 AA;  35797 MW;  6ABA46AEEB31A69A CRC64;

  Query Match             94.9%;  Score 37;  DB 2;  Length 320;
  Best Local Similarity   85.7%;  Pred. No. 2.2e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||:||||
Db        281 RPMPSRP 287


RESULT 7
```

```
Q9ABX1_CAUCR
ID   Q9ABX1_CAUCR PRELIMINARY;      PRT;   103 AA.
AC   Q9ABX1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein CC0099.
GN   OrderedLocusNames=CC0099;
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA   Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA   Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA   Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR   EMBL; AE005684; AAK22086.1; -; Genomic_DNA.
DR   PIR; B87261; B87261.
DR   TIGR; CC0099; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   103 AA;  11385 MW;  3EBA6CA59F7C5166 CRC64;

  Query Match             92.3%;  Score 36;  DB 2;  Length 103;
  Best Local Similarity   85.7%;  Pred. No. 90;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db         10 RPLPNRP 16


RESULT 8
Q4HB55_9DEIO
ID   Q4HB55_9DEIO PRELIMINARY;      PRT;   168 AA.
AC   Q4HB55;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=DgeoDRAFT_1858;
OS   Deinococcus geothermalis DSM 11300.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC   Deinococcaceae; Deinococcus.
OX   NCBI_TaxID=319795;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=DSM 11300;
RG   US DOE Joint Genome Institute (JGI-PGF);
RA   Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA   Hammon N., Israni S., Pitluck S., Richardson P.;
RT   "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT   DSM 11300.";
RL   Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=DSM 11300;
RG   US DOE Joint Genome Institute (JGI-ORNL);
RA   Larimer F., Land M.;
RT   "Annotation of the draft genome assembly of Deinococcus geothermalis
RT   DSM 11300.";
RL   Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
```

```
CC       preliminary data.
DR   EMBL; AAHE01000002; EAL83752.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   168 AA;  18848 MW;  BB7FF7CCC6435854 CRC64;

  Query Match             92.3%;  Score 36;  DB 2;  Length 168;
  Best Local Similarity   85.7%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db          7 RPLPTRP 13


RESULT 9
Q5SIN6_THET8
ID   Q5SIN6_THET8 PRELIMINARY;      PRT;   208 AA.
AC   Q5SIN6;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein TTHA1333.
GN   OrderedLocusNames=TTHA1333;
OS   Thermus thermophilus (strain HB8 / ATCC 27634 / DSM 579).
OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC   Thermus.
OX   NCBI_TaxID=300852;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=HB8;
RA   Masui R., Kurokawa K., Nakagawa N., Tokunaga F., Koyama Y.,
RA   Shibata T., Oshima T., Yokoyama S., Yasunaga T., Kuramitsu S.;
RT   "Complete genome sequence of Thermus thermophilus HB8.";
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP008226; BAD71156.1; -; Genomic_DNA.
DR   InterPro; IPR000637; A+T_hook.
DR   PRINTS; PR00929; ATHOOK.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   208 AA;  22208 MW;  FFFD38C58DB191BB CRC64;

  Query Match             92.3%;  Score 36;  DB 2;  Length 208;
  Best Local Similarity   85.7%;  Pred. No. 2e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db        109 RPLPARP 115


RESULT 10
Q9A584_CAUCR
ID   Q9A584_CAUCR PRELIMINARY;      PRT;   270 AA.
AC   Q9A584;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcriptional regulator, AraC family.
GN   OrderedLocusNames=CC2573;
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA   Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA   Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA   Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR   EMBL; AE005925; AAK24543.1; -; Genomic_DNA.
DR   PIR; C87568; C87568.
DR   TIGR; CC2573; -.
DR   GO; GO:0005622; C:intracellular; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR012287; Homeodomain-rel.
DR   InterPro; IPR000005; HTHAraC.
DR   Pfam; PF00165; HTH_AraC; 2.
DR   PRINTS; PR00032; HTHARAC.
DR   SMART; SM00342; HTH_ARAC; 1.
DR   PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW   Activator; Complete proteome; DNA-binding; Transcription;
KW   Transcription regulation.
SQ   SEQUENCE   270 AA;  30023 MW;  7A4700CF5FA37738 CRC64;

  Query Match             92.3%;  Score 36;  DB 2;  Length 270;
  Best Local Similarity   85.7%;  Pred. No. 2.7e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db         34 RPLPARP 40


RESULT 11
Q5NTG0_9BACT
ID   Q5NTG0_9BACT PRELIMINARY;      PRT;   313 AA.
AC   Q5NTG0;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hydrogen-peroxide-inducible genes activator.
GN   Name=bzo32-4;
OS   uncultured bacterium.
OC   Bacteria; environmental samples.
OX   NCBI_TaxID=77133;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15608629; DOI=10.1038/nbt1048;
RA   Uchiyama T., Abe T., Ikemura T., Watanabe K.;
RT   "Substrate-induced gene-expression screening of environmental
RT   metagenome libraries for isolation of catabolic genes.";
RL   Nat. Biotechnol. 23:88-93(2005).
CC   -!- SIMILARITY: Contains 1 HTH lysR-type DNA-binding domain.
DR   EMBL; AB190318; BAD81009.1; -; Genomic_DNA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   InterPro; IPR000847; HTH_LysR.
DR   InterPro; IPR005119; LysR_subst.
DR   Pfam; PF00126; HTH_1; 1.
DR   Pfam; PF03466; LysR_substrate; 1.
DR   PRINTS; PR00039; HTHLYSR.
DR   PROSITE; PS50931; HTH_LYSR; 1.
KW   DNA-binding; Transcription; Transcription regulation.
SQ   SEQUENCE   313 AA;  34191 MW;  DF336088553D3092 CRC64;

  Query Match             92.3%;  Score 36;  DB 2;  Length 313;
  Best Local Similarity   85.7%;  Pred. No. 3.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db        271 RPLPNRP 277


RESULT 12
Q8PLV6_XANAC
ID   Q8PLV6_XANAC PRELIMINARY;      PRT;   334 AA.
```

```
AC   Q8PLV6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein XAC1683.
GN   OrderedLocusNames=XAC1683;
OS   Xanthomonas axonopodis (pv. citri).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=92829;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=306 / ATCC 13902 / XV 101;
RX   MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA   Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA   Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA   Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA   El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA   Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE011800; AAM36550.1; -; Genomic_DNA.
DR   InterPro; IPR010239; Cons_hypoth2001.
DR   TIGRFAMs; TIGR02001; gcw_chp; 1.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   334 AA;  36710 MW;  C7449C413BBFB116 CRC64;

  Query Match           92.3%;  Score 36;  DB 2;  Length 334;
  Best Local Similarity   85.7%;  Pred. No. 3.4e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              :||||||
Db         49 KPLPSRP 55


RESULT 13
Q84SM1_ORYSA
ID   Q84SM1_ORYSA PRELIMINARY;      PRT;   352 AA.
AC   Q84SM1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein OJ1092_A07.118.
GN   Name=OJ1092_A07.118;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
RT   clone:OJ1092_A07.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP003866; BAC55662.1; -; Genomic_DNA.
DR   Gramene; Q84SM1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   352 AA;  38261 MW;  56890C8BBCE0C5F2 CRC64;

  Query Match           92.3%;  Score 36;  DB 2;  Length 352;
  Best Local Similarity   85.7%;  Pred. No. 3.6e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db        148 RPLPTRP 154


RESULT 14
Q8YXQ8_ANASP
ID   Q8YXQ8_ANASP PRELIMINARY;      PRT;   369 AA.
AC   Q8YXQ8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   All1154 protein.
GN   OrderedLocusNames=all1154;
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
DR   EMBL; BA000019; BAB73111.1; -; Genomic_DNA.
DR   PIR; AG1950; AG1950.
KW   Complete proteome.
SQ   SEQUENCE   369 AA;  39763 MW;  42DED3CB4EB8A922 CRC64;

  Query Match           92.3%;  Score 36;  DB 2;  Length 369;
  Best Local Similarity   85.7%;  Pred. No. 3.8e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              ||||:||
Db        168 RPLPTRP 174


RESULT 15
Q72GL1_THET2
ID   Q72GL1_THET2 PRELIMINARY;      PRT;   479 AA.
AC   Q72GL1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical membrane spanning protein.
GN   OrderedLocusNames=TTC1837;
OS   Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC   Thermus.
OX   NCBI_TaxID=262724;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15064768; DOI=10.1038/nbt956;
RA   Henne A., Brueggemann H., Raasch C., Wiezer A., Hartsch T.,
RA   Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA   Jacobi C., Starkuviene V., Schlenczeck S., Dencker S., Huber R.,
RA   Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT   "The genome sequence of the extreme thermophile Thermus
RT   thermophilus.";
RL   Nat. Biotechnol. 22:547-553(2004).
DR   EMBL; AE017307; AAS82179.1; -; Genomic_DNA.
DR   InterPro; IPR002110; ANK.
DR   InterPro; IPR007016; Wzy_C.
DR   Pfam; PF04932; Wzy_C; 1.
DR   PRINTS; PR01415; ANKYRIN.
```

```
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   479 AA;  51484 MW;  449ECB340F92C790 CRC64;

  Query Match             92.3%;  Score 36;  DB 2;  Length 479;
  Best Local Similarity   85.7%;  Pred. No. 5.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RPLPSRP 7
              :||||||
Db          4 KPLPSRP 10


Search completed: April  4, 2006, 13:15:22
Job time : 7.35079 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 3.47251 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-292
Perfect score: 38
Sequence: 1 SRLPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	7	3	AAB17236	Aab17236 SH3 antag
2	38	100.0	7	5	ABB73229	Abb73229 Src homol
3	38	100.0	7	7	ADJ73383	Adj73383 SH3 antag
4	38	100.0	7	8	ADJ53017	Adj53017 CH1 delet
5	38	100.0	7	8	ADJ51978	Adj51978 CH1 delet
6	38	100.0	13	2	AAW11115	Aaw11115 Src SH3 d
7	38	100.0	31	2	AAW16930	Aaw16930 Random re
8	38	100.0	31	2	AAW25493	Aaw25493 Random pe
9	38	100.0	61	7	ADA07698	Ada07698 Human sec
10	38	100.0	61	8	ADN41508	Adn41508 Novel hum
11	38	100.0	118	5	ADK34724	Adk34724 Novel hum
12	38	100.0	220	2	AAW63685	Aaw63685 Human sec
13	38	100.0	220	2	AAY03241	Aay03241 Clone HP1
14	38	100.0	220	2	AAY03240	Aay03240 Clone HP1
15	38	100.0	220	4	AAU29194	Aau29194 Human PRO
16	38	100.0	220	5	AAU83690	Aau83690 Human PRO
17	38	100.0	220	5	ABG79658	Abg79658 Invertebr
18	38	100.0	220	5	ADY31938	Ady31938 Novel hum
19	38	100.0	220	6	ABU58570	Abu58570 Human PRO
20	38	100.0	220	6	ABU88118	Abu88118 Novel hum
21	38	100.0	220	6	ABU84433	Abu84433 Human sec
22	38	100.0	220	6	ABR66307	Abr66307 Human sec
23	38	100.0	220	6	ABR65697	Abr65697 Human sec
24	38	100.0	220	6	ABU99637	Abu99637 Human sec
25	38	100.0	220	6	ABU82876	Abu82876 Human PRO
26	38	100.0	220	6	ABU89997	Abu89997 Novel hum
27	38	100.0	220	6	ABR68246	Abr68246 Human sec
28	38	100.0	220	6	ABU96299	Abu96299 Novel hum
29	38	100.0	220	6	ABU92730	Abu92730 Human sec
30	38	100.0	220	6	ABU80837	Abu80837 Human PRO
31	38	100.0	220	6	ABO08807	Abo08807 Human sec
32	38	100.0	220	6	ABO02859	Abo02859 Human sec
33	38	100.0	220	6	ABR75013	Abr75013 Human sec
34	38	100.0	220	6	ABR94775	Abr94775 Human sec
35	38	100.0	220	6	ABO33803	Abo33803 Novel hum
36	38	100.0	220	6	ABU85748	Abu85748 Human PRO
37	38	100.0	220	6	ABU98908	Abu98908 Novel hum
38	38	100.0	220	6	ABU98123	Abu98123 Novel hum
39	38	100.0	220	6	ABU91829	Abu91829 Novel hum
40	38	100.0	220	6	ABU89522	Abu89522 Human PRO
41	38	100.0	220	6	ABU86363	Abu86363 Human sec
42	38	100.0	220	6	ABU67576	Abu67576 Human sec
43	38	100.0	220	6	ABU80604	Abu80604 Human PRO
44	38	100.0	220	6	ABR99522	Abr99522 Human sec
45	38	100.0	220	6	ABR98912	Abr98912 Human sec

ALIGNMENTS

RESULT 1
AAB17236
ID AAB17236 standard; peptide; 7 AA.
XX
AC AAB17236;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:292.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 298; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRLPPLP 7
|||||||
Db 1 SRLPPLP 7

RESULT 2
ABB73229
ID ABB73229 standard; peptide; 7 AA.
XX
AC ABB73229;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:292.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRLPPLP 7
|||||||
Db 1 SRLPPLP 7

RESULT 3
ADJ73383
ID ADJ73383 standard; peptide; 7 AA.
XX
AC ADJ73383;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 838.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 838; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRLPPLP 7
|||||||
Db 1 SRLPPLP 7

RESULT 4
ADJ53017
ID ADJ53017 standard; peptide; 7 AA.
XX
AC ADJ53017;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID838.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 838; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRLPPLP 7
|||||||
Db 1 SRLPPLP 7

RESULT 5
ADJ51978
ID ADJ51978 standard; peptide; 7 AA.
XX
AC ADJ51978;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID838.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 838; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be

```
CC   useful for the development of compounds with an osteopathic,
CC   cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC   gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC   immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC   antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC   respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC   modulator or cytokine-agonist. The methods and compositions of the
CC   present invention are useful for the diagnosis, prevention and/or
CC   treatment of diseases or conditions associated with aberrant expression
CC   or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC   cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC   endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC   obstetric, haematologic, immunological, allergic, infectious,
CC   musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC   pediatric, psychiatric, renal or pulmonary disorders. The present
CC   sequence is that of a peptide which may be used during the creation of a
CC   mimetibody of the invention.
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 38;  DB 8;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db          1 SRLPPLP 7


RESULT 6
AAW11115
ID   AAW11115 standard; peptide; 13 AA.
XX
AC   AAW11115;
XX
DT   25-JUN-1997  (first entry)
XX
DE   Src SH3 domain-binding peptide used in signal transduction modulation.
XX
KW   Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW   protein tyrosine kinase; signal transduction; RNA processing;
KW   trafficking; translation.
XX
OS   Synthetic.
XX
PN   WO9603649-A1.
XX
PD   08-FEB-1996.
XX
PF   24-JUL-1995;   95WO-US009382.
XX
PR   22-JUL-1994;   94US-00278865.
PR   07-JUN-1995;   95US-00483555.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ;
XX
DR   WPI; 1996-117151/12.
XX
PT   Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT   proteins - useful for e.g. modulating signal transduction pathways at the
PT   cellular level, esp. protein tyrosine kinase-mediated.
XX
PS   Claim 39; Page 83; 116pp; English.
XX
CC   AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC   binding peptides are useful in modulating signal transduction pathways at
CC   the cellular level (especially protein tyrosine kinase-mediated),
CC   modulating oncogenic protein activity, or providing compounds for the
CC   development of drugs with the ability to modulate broad classes, as well
CC   as specific classes, of proteins involved in signal transduction and also
CC   for regulating the processing, trafficking or translation of RNA.
CC   Conjugates of the peptides with detectable labels or imaging agents are
CC   useful for imaging cells, tissues and organs in which Src or Src-related
CC   proteins are expressed
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 38;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db          4 SRLPPLP 10


RESULT 7
AAW16930
ID   AAW16930 standard; peptide; 31 AA.
XX
AC   AAW16930;
XX
DT   27-JUN-1997  (first entry)
XX
DE   Random recombinant SH3 domain binding peptide.
XX
KW   Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW   protein tyrosine kinase; signal transduction; RNA processing;
KW   trafficking; translation.
XX
OS   Synthetic.
XX
PN   WO9603649-A1.
XX
PD   08-FEB-1996.
XX
PF   24-JUL-1995;   95WO-US009382.
XX
PR   22-JUL-1994;   94US-00278865.
PR   07-JUN-1995;   95US-00483555.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ;
XX
DR   WPI; 1996-117151/12.
XX
PT   Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT   proteins - useful for e.g. modulating signal transduction pathways at the
PT   cellular level, esp. protein tyrosine kinase-mediated.
XX
PS   Disclosure; Fig 1; 116pp; English.
XX
CC   AAW16924-W16948 are random recombinant peptides derived from one of three
CC   peptide libraries, T9, T12 and R8C. The peptides are all SH3 domain-
CC   binding peptides. SH3 binding peptides are useful in modulating signal
CC   transduction pathways at the cellular level (especially protein tyrosine
CC   kinase-mediated), modulating oncogenic protein activity, or providing
CC   compounds for the development of drugs with the ability to modulate broad
CC   classes, as well as specific classes, of proteins involved in signal
CC   transduction and also for regulating the processing, trafficking or
CC   translation of RNA. Conjugates of the peptides with detectable labels or
CC   imaging agents are useful for imaging cells, tissues and organs in which
CC   Src or Src-related proteins are expressed
XX
SQ   Sequence 31 AA;

  Query Match             100.0%;  Score 38;  DB 2;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         19 SRLPPLP 25
```

```
RESULT 8
AAW25493
ID   AAW25493 standard; peptide; 31 AA.
XX
AC   AAW25493;
XX
DT   27-MAR-1998  (first entry)
XX
DE   Random peptide recombinant clone T12.SRC3.5.
XX
KW   Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW   tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW   PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS   Synthetic.
OS   Unidentified.
XX
PN   WO9730074-A1.
XX
PD   21-AUG-1997.
XX
PF   14-FEB-1997;   97WO-US002298.
XX
PR   16-FEB-1996;   96US-00602999.
XX
PA   (CYTO-) CYTOGEN CORP.
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ,  Fowlkes DM;
PI   Rider JE;
XX
DR   WPI; 1997-424972/39.
XX
PT   Src homology region 3 binding peptide - used to activate Src tyrosine
PT   kinase(s) and to stimulate immune response by increasing production of
PT   certain lymphokine(s), e.g. interleukin-1.
XX
PS   Disclosure; Fig 5; 131pp; English.
XX
CC   The present sequence represents a random peptide recombinant isolated by
CC   the method of the present invention. SH3 (Src homology region 3) binding
CC   peptides are selected from: (a) peptides which bind the SH3 domain of
CC   Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)
CC   peptides which bind the SH3 domain of Abl; (d) peptides which bind the
CC   SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;
CC   (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind
CC   the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3
CC   domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain
CC   of Grb2. The purified binding peptides can be used in the method to
CC   identify inhibitors of their binding to their respective SH3 domains,
CC   which could be used to modulate the pharmacological activity of proteins
CC   or polypeptide containing the SH3 domain. The peptides can also be used
CC   to activate Src or Src-related protein tyrosine kinases, to stimulate the
CC   immune response by increasing the production of certain lymphokines, e.g.
CC   tumour necrosis factor-alpha and interleukin-1, or to deliver a
CC   conjugated molecule to certain cellular compartments containing Src or
CC   Src related proteins
XX
SQ   Sequence 31 AA;

  Query Match             100.0%;  Score 38;  DB 2;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         19 SRLPPLP 25


RESULT 9
ADA07698
ID   ADA07698 standard; peptide; 61 AA.
XX
AC   ADA07698;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human secreted protein from gene 78, peptide #2.
XX
KW   Immunosuppressive; dermatological; antiinflammatory; antiallergic;
KW   antiarthritic; human; autoimmune disease; autoimmune disorder; lupus;
KW   transplant rejection; allergic reaction; arthritis;
KW   squamous cell E48 antigen.
XX
OS   Homo sapiens.
XX
PN   US2003064412-A1.
XX
PD   03-APR-2003.
XX
PF   30-OCT-2001; 2001US-00984490.
XX
PR   08-JUL-1997;   97US-0051916P.
PR   08-JUL-1997;   97US-0051918P.
PR   08-JUL-1997;   97US-0051919P.
PR   08-JUL-1997;   97US-0051920P.
PR   08-JUL-1997;   97US-0051925P.
PR   08-JUL-1997;   97US-0051926P.
PR   08-JUL-1997;   97US-0051928P.
PR   08-JUL-1997;   97US-0051929P.
PR   08-JUL-1997;   97US-0051930P.
PR   08-JUL-1997;   97US-0051931P.
PR   08-JUL-1997;   97US-0051932P.
PR   08-JUL-1997;   97US-0052732P.
PR   08-JUL-1997;   97US-0052733P.
PR   08-JUL-1997;   97US-0052793P.
PR   08-JUL-1997;   97US-0052795P.
PR   08-JUL-1997;   97US-0052803P.
PR   18-AUG-1997;   97US-0055684P.
PR   18-AUG-1997;   97US-0055722P.
PR   18-AUG-1997;   97US-0055723P.
PR   18-AUG-1997;   97US-0055947P.
PR   18-AUG-1997;   97US-0055948P.
PR   18-AUG-1997;   97US-0055949P.
PR   18-AUG-1997;   97US-0055950P.
PR   18-AUG-1997;   97US-0055953P.
PR   18-AUG-1997;   97US-0055954P.
PR   18-AUG-1997;   97US-0055964P.
PR   18-AUG-1997;   97US-0055984P.
PR   18-AUG-1997;   97US-0056360P.
PR   12-SEP-1997;   97US-0058660P.
PR   12-SEP-1997;   97US-0058661P.
PR   12-SEP-1997;   97US-0058664P.
PR   12-SEP-1997;   97US-0058785P.
PR   07-JUL-1998;   98WO-US013684.
PR   08-JAN-1999;   99US-00227357.
XX
PA   (FISC/) FISCHER C L.
PA   (ROSE/) ROSEN C A.
PA   (SOPP/) SOPPET D R.
PA   (RUBE/) RUBEN S M.
PA   (KYAW/) KYAW H.
PA   (LIYY/) LI Y.
PA   (ZENG/) ZENG Z.
PA   (LAFL/) LAFLEUR D W.
PA   (MOOR/) MOORE P A.
PA   (SHIY/) SHI Y.
PA   (OLSE/) OLSEN H S.
PA   (EBNE/) EBNER R.
PA   (BREW/) BREWER L A.
XX
PI   Fischer CL,  Rosen CA,  Soppet DR,  Ruben SM,  Kyaw H,  Li Y,  Zeng Z;
PI   Lafleur DW,  Moore PA,  Shi Y,  Olsen HS,  Ebner R,  Brewer LA;
XX
```

DR WPI; 2003-540785/51.
XX
PT Novel antibody which specifically binds to a secreted protein useful for
PT diagnosing and treating lupus, arthritis, allergic reactions, arthritis.
XX
PS Disclosure; Page 54; 355pp; English.
XX
CC The invention relates to an isolated antibody or its portion that
CC specifically binds to a protein that shares sequence homology with human
CC squamous cell E48 antigen, and consists of amino acid residues 21-116 or
CC 1-116 the protein appearing as ADA07417 (one of 123 disclosed novel human
CC secreted proteins encoded by 123 novel genes), or a protein consisting of
CC amino acid sequence of secreted or full-length polypeptide encoded by
CC HLHFPO3 cDNA contained in ATCC Deposit No. 209126. The antibody is
CC produced by immunising an animal with amino acid residues 21-116 of
CC ADA07417, or with a protein consisting of amino acid sequence of the
CC secreted polypeptide encoded by the HLHFPO3 cDNA contained in ATCC
CC Deposit No. 209126, respectively. Also included are an isolated cell that
CC produces the antibody and a hybridoma that produces the antibody. The
CC antibody is a monoclonal, polyclonal, chimeric, humanised or human
CC antibody. Optionally, the antibody is a Fab fragment, and is labelled by
CC a label chosen from enzyme label, a radioisotope, and a fluorescent
CC label. The antibody is useful as a probe for differential identification
CC of tissues or cell types in which ADA07417 is expressed. The antibody is
CC also for diagnosis and treatment of autoimmune diseases and disorders,
CC such as lupus, transplant rejection, allergic reactions, and arthritis.
CC The present sequence is a peptide/protein derived from one of the 123
CC novel secreted proteins.
XX
SQ Sequence 61 AA;

Query Match 100.0%; Score 38; DB 7; Length 61;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRLPPLP 7
|||||||
Db 4 SRLPPLP 10

RESULT 10
ADN41508
ID ADN41508 standard; protein; 61 AA.
XX
AC ADN41508;
XX
DT 17-JUN-2004 (first entry)
XX
DE Novel human secreted protein fragment seqid 630.
XX
KW immunomodulator; immunosuppressive; antiinflammatory; dermatological;
KW antiarthritic; antirheumatic; neuroprotective; antianaemic; muscular;
KW antiallergic; antiasthmatic; gastrointestinal; anticoagulant;
KW thrombolytic; antiarteriosclerotic; cardiant; cytostatic; nephrotropic;
KW cardiovascular; respiratory; gene therapy; secreted protein;
KW chromosome identification; hybrid mapping; gene expression control;
KW immune system disorder; immunodeficiency; Chediak-Higashi syndrome;
KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
KW multiple sclerosis; haemolytic anaemia; myasthenia gravis;
KW allergic reaction; asthma; inflammatory condition;
KW inflammatory bowel disease; B cell stimulator; T cell activator;
KW blood-related disorder; eosinophilia; thrombosis; thromboembolism;
KW atherosclerosis; myocardial infarction; angina; anaemia;
KW hyperproliferative disorder; cancer; renal disorder;
KW chronic kidney failure; renal tubular acidosis; kidney stone;
KW cardiovascular disorder; respiratory disorder; human.
XX
OS Homo sapiens.
XX
PN US2004044191-A1.
XX
PD 04-MAR-2004.
XX
PF 10-OCT-2001; 2001US-00973278.
XX
PR 08-JUL-1997; 97US-0051916P.
PR 08-JUL-1997; 97US-0051918P.
PR 08-JUL-1997; 97US-0051919P.
PR 08-JUL-1997; 97US-0051920P.
PR 08-JUL-1997; 97US-0051925P.
PR 08-JUL-1997; 97US-0051926P.
PR 08-JUL-1997; 97US-0051928P.
PR 08-JUL-1997; 97US-0051929P.
PR 08-JUL-1997; 97US-0051930P.
PR 08-JUL-1997; 97US-0051931P.
PR 08-JUL-1997; 97US-0051932P.
PR 08-JUL-1997; 97US-0052732P.
PR 08-JUL-1997; 97US-0052733P.
PR 08-JUL-1997; 97US-0052793P.
PR 08-JUL-1997; 97US-0052795P.
PR 08-JUL-1997; 97US-0052803P.
PR 18-AUG-1997; 97US-0055684P.
PR 18-AUG-1997; 97US-0055722P.
PR 18-AUG-1997; 97US-0055723P.
PR 18-AUG-1997; 97US-0055947P.
PR 18-AUG-1997; 97US-0055948P.
PR 18-AUG-1997; 97US-0055949P.
PR 18-AUG-1997; 97US-0055950P.
PR 18-AUG-1997; 97US-0055953P.
PR 18-AUG-1997; 97US-0055954P.
PR 18-AUG-1997; 97US-0055964P.
PR 18-AUG-1997; 97US-0055984P.
PR 18-AUG-1997; 97US-0056360P.
PR 12-SEP-1997; 97US-0058660P.
PR 12-SEP-1997; 97US-0058661P.
PR 12-SEP-1997; 97US-0058664P.
PR 12-SEP-1997; 97US-0058785P.
PR 07-JUL-1998; 98WO-US013684.
PR 08-JAN-1999; 99US-00227357.
PR 13-OCT-2000; 2000US-0239899P.
XX
PA (FISC/) FISCHER C L.
PA (ROSE/) ROSEN C A.
PA (SOPP/) SOPPET D R.
PA (RUBE/) RUBEN S M.
PA (KYAW/) KYAW H.
PA (LIYY/) LI Y.
PA (ZENG/) ZENG Z.
PA (LAFL/) LAFLEUR D W.
PA (MOOR/) MOORE P A.
PA (SHIY/) SHI Y.
PA (OLSE/) OLSEN H.
PA (EBNE/) EBNER R.
PA (BIRS/) BIRSE C E.
XX
PI Fischer CL, Rosen CA, Soppet DR, Ruben SM, Kyaw H, Li Y, Zeng Z;
PI Lafleur DW, Moore PA, Shi Y, Olsen H, Ebner R, Birse CE;
XX
DR WPI; 2004-225733/21.
XX
PT New isolated nucleic acid encoding human proteins, useful for treating,
PT preventing or diagnosing e.g. rheumatoid arthritis, multiple sclerosis,
PT anemia, inflammatory bowel disease, atherosclerosis, cancers, chronic
PT kidney failure.
XX
PS Disclosure; SEQ ID NO 630; 372pp; English.
XX
CC The invention describes novel human secreted proteins and the nucleotides
CC encoding them. The polynucleotides are useful in chromosome
CC identification, for radiation hybrid mapping, in controlling gene
CC expression, in gene therapy or as molecular weight markers. The
CC polynucleotides and polypeptides are useful for diagnosing, treating or
CC preventing diseases of the immune system, immunodeficiencies, e.g.
CC Chediak-Higashi syndrome, autoimmune diseases, e.g. systemic lupus
CC erythematosus, rheumatoid arthritis, multiple sclerosis, haemolytic
CC anaemia or myasthenia gravis, allergic reactions, e.g. asthma,

```
CC   inflammatory conditions, e.g. inflammatory bowel disease. They can also
CC   be used as a stimulator of B cell responsiveness to pathogens or as an
CC   activator of T cells. The polynucleotides and polypeptides are also
CC   useful for treating or preventing blood-related disorders, e.g.
CC   eosinophilia, thrombosis, thromboembolism, atherosclerosis, myocardial
CC   infarction, unstable angina or anaemia. They can also be used for
CC   treating, preventing or diagnosing hyperproliferative disorders
CC   (cancers), renal disorders (chronic kidney failure, renal tubular
CC   acidosis or kidney stones), cardiovascular disorders or respiratory
CC   disorders. This is the amino acid sequence of a novel human secreted
CC   protein fragment. Note: This sequence is available in electronic format
CC   from the US patent office at
CC   ftp.seqdata.uspto.gov/sequence.html?DocID=20040044191.
XX
SQ   Sequence 61 AA;

  Query Match             100.0%;  Score 38;  DB 8;  Length 61;
  Best Local Similarity   100.0%;  Pred. No. 82;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db          4 SRLPPLP 10


RESULT 11
ADK34724
ID   ADK34724 standard; protein; 118 AA.
XX
AC   ADK34724;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Novel human polypeptide SeqID6806.
XX
KW   antiarthritic; antiparkinsonian; neuroprotective; nootropic;
KW   immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
KW   antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
KW   arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
KW   psoriasis; inflammatory bowel disease; infection; bacteria; virus;
KW   fungus; parasite; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .118
FT                   /label= OTHER
FT                   /note= "OTHER= All Xaa's in this sequence are unknown
FT                   amino acids or the site of a stop codon within the DNA
FT                   sequence"
XX
PN   WO200216439-A2.
XX
PD   28-FEB-2002.
XX
PF   05-MAR-2001; 2001WO-US004941.
XX
PR   07-MAR-2000; 2000US-00519705.
PR   19-MAY-2000; 2000US-00574454.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2002-280918/32.
XX
PT   Isolated polynucleotide encoding bone marrow derived polypeptides useful
PT   for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
PT   disease, and inflammatory bowel disease.
XX
PS   Claim 20; SEQ ID NO 6806; 504pp; English.
XX
CC   This invention relates to a novel isolated polynucleotide comprising a
CC   nucleotide sequence selected from one of 1680 sequences, a mature protein
CC   coding portion of them, an active domain of them and their complementary
CC   sequences. The invention may be useful for the production of compounds
CC   with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
CC   immunosuppressive, cytostatic, antipsoriatic, antiinflammatory,
CC   antibacterial, antiviral, antifungal or antiparasitic activity. In
CC   addition, the disclosed sequences may be useful for gene therapy. The
CC   polypeptides or their antibodies are useful for treating many diseases
CC   such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
CC   psoriasis, inflammatory bowel disease and infections caused by bacteria,
CC   viruses, fungi or parasites. The present sequence is that of a human
CC   polypeptide of the invention.
XX
SQ   Sequence 118 AA;

  Query Match             100.0%;  Score 38;  DB 5;  Length 118;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         60 SRLPPLP 66


RESULT 12
AAW63685
ID   AAW63685 standard; protein; 220 AA.
XX
AC   AAW63685;
XX
DT   24-SEP-1998  (first entry)
XX
DE   Human secreted protein 5.
XX
KW   Secreted protein; human; cell proliferation; cytokine activity;
KW   tissue growth; cellular differentiation; regeneration; activin; inhibin;
KW   chemotactic; haemostatic; thrombolytic; tumour inhibition;
KW   anti-inflammatory activity; biomarker.
XX
OS   Homo sapiens.
XX
PN   WO9825959-A2.
XX
PD   18-JUN-1998.
XX
PF   11-DEC-1997;   97WO-US022787.
XX
PR   11-DEC-1996;   96US-0032757P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Escobedo J,  Hu Q,  Garcia P,  Williams LT,  Kothakota S;
XX
DR   WPI; 1998-348453/30.
DR   N-PSDB; AAV43605.
XX
PT   Secreted human polypeptides - having cytokine, cell proliferation or
PT   differentiation, activin or inhibin, tumour inhibition or anti-
PT   inflammatory activities.
XX
PS   Claim 1; Page 53; 78pp; English.
XX
CC   This represents a human secreted protein. The specification provides
CC   secreted protein sequences (AAW63681 to AAW63699) encoded by the nucleic
CC   acid sequences shown in AAV43601 to AAV43619. The invention provides a
CC   method of identifying a secreted polypeptide which is modified by rough
CC   microsomes. The secreted proteins can be used in assays to determine
CC   biological activities, such as cytokine, cell proliferation, or cellular
CC   differentiation activities, tissue growth or regeneration, activin or
CC   inhibin activity, chemotactic or chemokinetic activity, haemostatic or
CC   thrombolytic activity, receptor/ligand activity, tumour inhibition, or
CC   anti-inflammatory activity. The proteins can also be used as biomarkers,
CC   to identify tissues or cell types which express the proteins, or a stage-
```

```
CC   or disease-specific alteration in protein expression. They can be used in
CC   protein interaction assays, to identify ligands or binding proteins.
CC   Compounds which affect the biological activities of the secreted proteins
CC   or their ability to interact with specific ligands can be identified
CC   using the proteins in screening assays. The proteins and antibodies that
CC   bind specifically to the protein can also be used to design diagnostic
CC   tests and therapeutic compositions for diseases which may be associated
CC   with altered expression of these proteins. Fusion proteins comprising,
CC   e.g. signal sequences or transmembrane domains of the proteins can be
CC   used to target other protein domains to cellular membrane or they can be
CC   secreted extracellularly
XX
SQ   Sequence 220 AA;

  Query Match             100.0%;  Score 38;  DB 2;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         44 SRLPPLP 50


RESULT 13
AAY03241
ID   AAY03241 standard; protein; 220 AA.
XX
AC   AAY03241;
XX
DT   26-AUG-1999  (first entry)
XX
DE   Clone HP10484 of a human secretory signal protein (2).
XX
KW   Human; secretory signal protein sequence; cell membrane; proliferation;
KW   differentiation; carcinostatic agent; antigen; antibody; probe;
KW   hybridisation; gene therapy; HP10484.
XX
OS   Homo sapiens.
XX
PN   WO9918204-A2.
XX
PD   15-APR-1999.
XX
PF   05-OCT-1998;   98WO-JP004476.
XX
PR   08-OCT-1997;   97JP-00276268.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Yamaguchi T,  Sekine S,  Kobayashi M;
XX
DR   WPI; 1999-264020/22.
DR   N-PSDB; AAX28686.
XX
PT   Human proteins with secretory signal sequences and nucleotide sequences.
XX
PS   Disclosure; Page 84; 84pp; English.
XX
CC   This is the amino acid sequence of a clone of a human secretory signal
CC   protein sequence, used in the method of the invention. All of the
CC   proteins exist in the cell membrane, so are considered to be proteins
CC   controlling the proliferation and differentiation of the cells. They may
CC   be useful as carcinostatic agents or as antigens for preparing antibodies
CC   against the proteins. The cDNAs can be used as probes for gene diagnosis
CC   and gene sources for gene therapy, as well as for large-scale expression
CC   of the proteins
XX
SQ   Sequence 220 AA;

  Query Match             100.0%;  Score 38;  DB 2;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         44 SRLPPLP 50


RESULT 14
AAY03240
ID   AAY03240 standard; protein; 220 AA.
XX
AC   AAY03240;
XX
DT   26-AUG-1999  (first entry)
XX
DE   Clone HP10484 of a human secretory signal protein (1).
XX
KW   Human; secretory signal protein sequence; cell membrane; proliferation;
KW   differentiation; carcinostatic agent; antigen; antibody; probe;
KW   hybridisation; gene therapy; HP10484.
XX
OS   Homo sapiens.
XX
PN   WO9918204-A2.
XX
PD   15-APR-1999.
XX
PF   05-OCT-1998;   98WO-JP004476.
XX
PR   08-OCT-1997;   97JP-00276268.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Yamaguchi T,  Sekine S,  Kobayashi M;
XX
DR   WPI; 1999-264020/22.
DR   N-PSDB; AAX28682.
XX
PT   Human proteins with secretory signal sequences and nucleotide sequences.
XX
PS   Claim 1; Page 69-70; 84pp; English.
XX
CC   This is the amino acid sequence of a clone of a human secretory signal
CC   protein sequence, used in the method of the invention. All of the
CC   proteins exist in the cell membrane, so are considered to be proteins
CC   controlling the proliferation and differentiation of the cells. They may
CC   be useful as carcinostatic agents or as antigens for preparing antibodies
CC   against the proteins. The cDNAs can be used as probes for gene diagnosis
CC   and gene sources for gene therapy, as well as for large-scale expression
CC   of the proteins
XX
SQ   Sequence 220 AA;

  Query Match             100.0%;  Score 38;  DB 2;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         44 SRLPPLP 50


RESULT 15
AAU29194
ID   AAU29194 standard; protein; 220 AA.
XX
AC   AAU29194;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human PRO polypeptide sequence #171.
XX
KW   PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
```

KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006520.
XX
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 03-MAR-2000; 2000US-0187202P.
PR 06-MAR-2000; 2000US-0186968P.
PR 14-MAR-2000; 2000US-0189320P.
PR 14-MAR-2000; 2000US-0189328P.
PR 15-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000US-0190828P.
PR 21-MAR-2000; 2000US-0191007P.
PR 21-MAR-2000; 2000US-0191048P.
PR 21-MAR-2000; 2000US-0191314P.
PR 28-MAR-2000; 2000US-0192655P.
PR 29-MAR-2000; 2000US-0193032P.
PR 29-MAR-2000; 2000US-0193053P.
PR 30-MAR-2000; 2000WO-US008439.
PR 04-APR-2000; 2000US-0194449P.
PR 04-APR-2000; 2000US-0194647P.
PR 11-APR-2000; 2000US-0195975P.
PR 11-APR-2000; 2000US-0196000P.
PR 11-APR-2000; 2000US-0196187P.
PR 11-APR-2000; 2000US-0196690P.
PR 11-APR-2000; 2000US-0196820P.
PR 18-APR-2000; 2000US-0198121P.
PR 18-APR-2000; 2000US-0198585P.
PR 25-APR-2000; 2000US-0199397P.
PR 25-APR-2000; 2000US-0199550P.
PR 25-APR-2000; 2000US-0199654P.
PR 03-MAY-2000; 2000US-0201516P.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 05-JUN-2000; 2000US-0209832P.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-602746/68.
DR N-PSDB; AAS46095.
XX
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumors, such as prostate and breast tumors, in mammals and to
PT screen for modulators of the compounds.
XX
PS Claim 11; Fig 342; 774pp; English.
XX
CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders
XX
SQ Sequence 220 AA;

Query Match 100.0%; Score 38; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRLPPLP 7
|||||||
Db 44 SRLPPLP 50

Search completed: April 4, 2006, 13:07:37
Job time : 5.47251 secs

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 1.14529 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-292
Perfect score: 38
Sequence: 1 SRLPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	313	2	T29195	hypothetical prote
2	38	100.0	360	2	S09792	hypothetical prote
3	38	100.0	657	2	T22451	hypothetical prote
4	38	100.0	781	2	T26080	hypothetical prote
5	36	94.7	249	2	A96632	hypothetical prote
6	35	92.1	262	2	A72469	hypothetical prote
7	35	92.1	789	2	S44759	C14B9.5 protein -
8	35	92.1	831	2	E70620	probable pheT prot
9	34	89.5	95	2	S77567	ribosomal protein
10	34	89.5	102	2	G25035	hypothetical prote
11	34	89.5	102	2	H25035	hypothetical prote
12	34	89.5	114	2	T45181	hypothetical prote
13	34	89.5	235	2	A83970	hypothetical prote
14	34	89.5	293	2	C83845	superoxide dismuta
15	34	89.5	295	2	E83058	hypothetical prote
16	34	89.5	405	2	T23321	hypothetical prote
17	34	89.5	445	2	A75376	probable oligosacc
18	34	89.5	731	2	B86369	hypothetical prote
19	34	89.5	832	2	T49494	condensin complex
20	34	89.5	1262	2	T25168	hypothetical prote
21	34	89.5	1590	2	B86398	protein T7N9.24 [i
22	33	86.8	340	1	WMBEL1	latency-related pr
23	33	86.8	355	2	T14086	hypothetical prote
24	33	86.8	478	2	A83368	hypothetical prote
25	33	86.8	558	2	G96522	F11A17.16 [importe
26	33	86.8	651	2	T31175	hypothetical prote
27	33	86.8	817	2	S53919	hypothetical prote
28	33	86.8	879	2	S49910	chloroplast outer
29	32	84.2	108	2	T51873	hypothetical prote
30	32	84.2	124	2	G87326	hypothetical prote
31	32	84.2	198	2	E75599	conserved hypothet
32	32	84.2	203	2	D81934	probable periplasm
33	32	84.2	203	2	F81171	cryptic protein NM
34	32	84.2	223	2	A23036	nodulin-23 - soybe
35	32	84.2	224	2	S07315	nodulin - soybean
36	32	84.2	246	2	F95397	Probable haloacid
37	32	84.2	284	2	G75447	hypothetical prote
38	32	84.2	361	2	A36669	3-alpha-galactosyl
39	32	84.2	432	2	A43448	thrombin receptor
40	32	84.2	443	2	A38219	GAP-associated tyr
41	32	84.2	447	2	T20249	hypothetical prote
42	32	84.2	452	2	F86289	probable cyclin [i
43	32	84.2	462	1	S35534	adenovirus E1A enh
44	32	84.2	525	2	T23304	hypothetical prote
45	32	84.2	534	2	T23305	hypothetical prote

ALIGNMENTS

RESULT 1
T29195
hypothetical protein T03F1.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004
C;Accession: T29195
R;Du, Z.; Le, T.T.
submitted to the EMBL Data Library, February 1997
A;Description: The sequence of C. elegans cosmid T03F1.
A;Reference number: Z20586
A;Accession: T29195
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-313 <DUZ>
A;Cross-references: UNIPROT:P91424; UNIPARC:UPI0000175301; EMBL:U88169; PIDN:AAB42235.
A;Experimental source: strain Bristol N2; clone T03F1
C;Genetics:
A;Gene: CESP:T03F1.7
A;Map position: 1
A;Introns: 41/3; 90/3; 153/2; 214/1
C;Superfamily: dimethyladenosine transferase (rRNA adenosine dimethyltransferase)

Query Match 100.0%; Score 38; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRLPPLP 7
|||||||
Db 5 SRLPPLP 11

RESULT 2
S09792
hypothetical protein UL29 - human cytomegalovirus (strain AD169)
C;Species: human cytomegalovirus, human herpesvirus 5
A;Note: host Homo sapiens (man)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S09792
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T. M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A;Title: Analysis of the protein-coding content of the sequence of human cytomegalovir
A;Reference number: S09749; MUID:90269039; PMID:2161319
A;Accession: S09792
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-360 <CHE>
A;Cross-references: UNIPROT:P16764; UNIPARC:UPI0000137B88; EMBL:X17403; NID:g59591; PI
A;Note: this sequence was submitted to the EMBL Data Library, December 1989

Query Match 100.0%; Score 38; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SRLPPLP 7
              |||||||
Db         22 SRLPPLP 28
```

RESULT 3
T22451
hypothetical protein F49E12.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22451
R;Thomas, K.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z19565
A;Accession: T22451
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-657 <WIL>
A;Cross-references: UNIPROT:Q20619; UNIPARC:UPI000007DBB4; EMBL:Z66520; PIDN:CAA91391.1;
A;Experimental source: clone F49E12
C;Genetics:
A;Gene: CESP:F49E12.6
A;Map position: 2
A;Introns: 30/3; 75/1; 133/1; 176/3; 276/3; 453/2; 590/2

Query Match 100.0%; Score 38; DB 2; Length 657;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SRLPPLP 7
              |||||||
Db        478 SRLPPLP 484
```

RESULT 4
T26080
hypothetical protein W02A2.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26080
R;Ainscough, R.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z20148
A;Accession: T26080
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-781 <WIL>
A;Cross-references: UNIPROT:Q9XUB3; UNIPARC:UPI000007A5B9; EMBL:Z82286; PIDN:CAB05309.1;
A;Experimental source: clone W02A2
C;Genetics:
A;Gene: CESP:W02A2.6
A;Map position: 4
A;Introns: 18/2; 85/1; 150/3; 189/2; 671/1; 720/2; 753/1

Query Match 100.0%; Score 38; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SRLPPLP 7
              |||||||
Db        354 SRLPPLP 360
```

RESULT 5
A96632
hypothetical protein F8A5.19 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A96632
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96632
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: UNIPROT:O22705; UNIPARC:UPI00000483C3; GB:AE005173; NID:g2462760; P
C;Genetics:
A;Gene: F8A5.19
A;Map position: 1

Query Match 94.7%; Score 36; DB 2; Length 249;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SRLPPLP 7
              ||:||||
Db        208 SRMPPLP 214
```

RESULT 6
A72469
hypothetical protein APE2394 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: A72469
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: A72469
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-262 <KAW>
A;Cross-references: UNIPROT:Q9Y991; UNIPARC:UPI000005E31D; DDBJ:AP000064; NID:g5105945;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2394
C;Superfamily: Aeropyrum pernix hypothetical protein APE2394

Query Match 92.1%; Score 35; DB 2; Length 262;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SRLPPLP 7
              ||:||||
Db        232 SRVPPLP 238
```

RESULT 7
S44759
C14B9.5 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Sep-1997
C;Accession: S44759
R;Favello, A.D.
submitted to the EMBL Data Library, May 1993
A;Description: Sequence of the C. elegans cosmid C14B9.
A;Reference number: S44617
A;Accession: S44759
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-789 <FAV>
A;Cross-references: UNIPARC:UPI000017B6AE; EMBL:L15188; NID:g289640; PID:g289646
C;Genetics:

```
A;Introns: 61/3; 129/2; 147/3; 191/3; 279/3; 368/3; 392/3; 627/3; 710/1; 731/1

  Query Match             92.1%;  Score 35;  DB 2;  Length 789;
  Best Local Similarity   85.7%;  Pred. No. 1.7e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              :||||||
Db        723 NRLPPLP 729


RESULT 8
E70620
probable pheT protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70620
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70620
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-831 <COL>
A;Cross-references: UNIPROT:P94985; UNIPARC:UPI0000136454; GB:Z85982; GB:AL123456; NID:g
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: pheT
C;Superfamily: phenylalanine-tRNA ligase beta chain

  Query Match             92.1%;  Score 35;  DB 2;  Length 831;
  Best Local Similarity   85.7%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              ||:||||
Db        201 SRVPPLP 207


RESULT 9
S77567
ribosomal protein S37, mitochondrial - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YDL045w-a; ribosomal protein YmS-T
C;Species: Saccharomyces cerevisiae
C;Date: 16-Apr-1997 #sequence_revision 16-Apr-1997 #text_change 09-Jul-2004
C;Accession: S77567; S78037
R;Jin, C.; Myers, A.M.; Tzagoloff, A.
Curr. Genet. 31, 228-234, 1997
A;Title: Cloning and characterization of MRP10 a yeast gene coding for a mitochondrial r
A;Reference number: S77567; MUID:97218168; PMID:9065385
A;Accession: S77567
A;Molecule type: DNA
A;Residues: 1-95 <JIN>
A;Cross-references: UNIPROT:O75012; UNIPARC:UPI000005308B; EMBL:Z71781
R;Kitakawa, M.; Graack, H.R.; Grohmann, L.; Goldschmidt-Reisin, S.; Herfurth, E.; Wittma
Eur. J. Biochem. 245, 449-456, 1997
A;Title: Identification and characterization of the genes for mitochondrial ribosomal pr
A;Reference number: S78018; MUID:97296414; PMID:9151978
A;Accession: S78037
A;Molecule type: protein
A;Residues: 'D',3-14,'XI',17-18,'X',20,'I' <KIT>
A;Cross-references: UNIPARC:UPI000017B33B
C;Genetics:
A;Gene: SGD:MRP10
A;Cross-references: SGD:S0006430; MIPS:YDL045w-a
A;Map position: 4L
A;Genome: nuclear
C;Keywords: mitochondrion; protein biosynthesis; ribosome

  Query Match             89.5%;  Score 34;  DB 2;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLPPLP 7
              ||||||
Db          9 RLPPLP 14


RESULT 10
G25035
hypothetical protein 2 - Escherichia coli plasmid ColIa
C;Species: Escherichia coli
C;Date: 24-Jan-1988 #sequence_revision 24-Jan-1988 #text_change 09-Jul-2004
C;Accession: G25035
R;Mankovich, J.A.; Hsu, C.H.; Konisky, J.
J. Bacteriol. 168, 228-236, 1986
A;Title: DNA and amino acid sequence analysis of structural and immunity genes of coli
A;Reference number: A91822; MUID:87008385; PMID:3531169
A;Accession: G25035
A;Molecule type: DNA
A;Residues: 1-102 <MAN>
A;Cross-references: UNIPROT:Q47295; UNIPARC:UPI00000B9259
C;Genetics:
A;Genome: plasmid

  Query Match             89.5%;  Score 34;  DB 2;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLPPLP 7
              ||||||
Db         25 RLPPLP 30


RESULT 11
H25035
hypothetical protein 2 - Escherichia coli plasmid ColIb
C;Species: Escherichia coli
C;Date: 24-Jan-1988 #sequence_revision 24-Jan-1988 #text_change 09-Jul-2004
C;Accession: H25035
R;Mankovich, J.A.; Hsu, C.H.; Konisky, J.
J. Bacteriol. 168, 228-236, 1986
A;Title: DNA and amino acid sequence analysis of structural and immunity genes of colic
A;Reference number: A91822; MUID:87008385; PMID:3531169
A;Accession: H25035
A;Molecule type: DNA
A;Residues: 1-102 <MAN>
A;Cross-references: UNIPROT:Q47298; UNIPARC:UPI00000B4296
C;Genetics:
A;Genome: plasmid

  Query Match             89.5%;  Score 34;  DB 2;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLPPLP 7
              ||||||
Db         25 RLPPLP 30


RESULT 12
T45181
hypothetical protein u1756m [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45181
R;Robison, K.
submitted to the EMBL Data Library, September 1994
A;Reference number: Z16911
A;Accession: T45181
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
```

A;Residues: 1-114 <KEI>
A;Cross-references: UNIPROT:Q49956; UNIPARC:UPI00000D4395; EMBL:U15180; PIDN:AAA62893.1
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1209

Query Match 89.5%; Score 34; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RLPPLP 7
              ||||||
Db         31 RLPPLP 36
```

RESULT 13
A83970
hypothetical protein BH2561 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: A83970
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: A83970
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-235 <STO>
A;Cross-references: UNIPROT:Q9K9T4; UNIPARC:UPI00000C3F39; GB:AP001515; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2561

Query Match 89.5%; Score 34; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RLPPLP 7
              ||||||
Db        169 RLPPLP 174
```

RESULT 14
C83845
superoxide dismutase BH1563 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: C83845
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: C83845
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-293 <STO>
A;Cross-references: UNIPROT:Q9KCK8; UNIPARC:UPI00000C3C25; GB:AP001512; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1563
C;Superfamily: superoxide dismutase (Mn)

Query Match 89.5%; Score 34; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RLPPLP 7
              ||||||
Db         93 RLPPLP 98
```

RESULT 15
E83058
hypothetical protein PA4705 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83058
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83058
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-295 <STO>
A;Cross-references: UNIPROT:Q9HV92; UNIPARC:UPI00000C5DD4; GB:AE004884; GB:AE004091; NI
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4705

Query Match 89.5%; Score 34; DB 2; Length 295;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 SRLPPLP 7
              |:|||||
Db         17 SQLPPLP 23
```

Search completed: April 4, 2006, 13:17:22
Job time : 3.14529 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 5.35079 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-292
Perfect score: 38
Sequence: 1 SRLPPLP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	120	2	Q5BB77_EMENI	Q5bb77 aspergillus
2	38	100.0	220	1	CREG1_HUMAN	O75629 homo sapien
3	38	100.0	318	2	Q57XH6_9TRYP	Q57xh6 trypanosoma
4	38	100.0	358	2	Q61ER4_CAEBR	Q61er4 caenorhabdi
5	38	100.0	360	1	UL29_HCMVA	P16764 human cytom
6	38	100.0	360	2	Q6XNK8_HCMV	Q6xnk8 human cytom
7	38	100.0	367	2	P91424_CAEEL	P91424 caenorhabdi
8	38	100.0	373	2	Q9H8W6_HUMAN	Q9h8w6 homo sapien
9	38	100.0	420	2	Q5RAU1_PONPY	Q5rau1 pongo pygma
10	38	100.0	428	2	O14498_HUMAN	O14498 homo sapien
11	38	100.0	428	2	Q5NVQ6_PONPY	Q5nvq6 pongo pygma
12	38	100.0	517	2	Q4R9C3_MACFA	Q4r9c3 macaca fasc
13	38	100.0	600	2	Q20619_CAEEL	Q20619 caenorhabdi
14	38	100.0	630	2	Q6NUN6_HUMAN	Q6nun6 homo sapien
15	38	100.0	654	2	Q68DD4_HUMAN	Q68dd4 homo sapien
16	38	100.0	654	2	Q9H6X0_HUMAN	Q9h6x0 homo sapien
17	38	100.0	686	2	Q8N384_HUMAN	Q8n384 homo sapien
18	38	100.0	694	2	Q8C7U9_MOUSE	Q8c7u9 mus musculu
19	38	100.0	694	2	Q6DI97_MOUSE	Q6di97 mus musculu
20	38	100.0	694	2	Q501J7_MOUSE	Q501j7 mus musculu
21	38	100.0	702	2	Q8IZ21_HUMAN	Q8iz21 homo sapien
22	38	100.0	706	2	Q5DTM4_MOUSE	Q5dtm4 mus musculu
23	38	100.0	781	2	Q9XUB3_CAEEL	Q9xub3 caenorhabdi
24	38	100.0	1178	2	Q7QR62_GIALA	Q7qr62 giardia lam
25	38	100.0	2371	2	Q5B0E8_EMENI	Q5b0e8 aspergillus
26	38	100.0	2385	2	Q6KBA5_EMENI	Q6kba5 emericella
27	36	94.7	208	2	Q5SKG2_THET8	Q5skg2 thermus the
28	36	94.7	208	2	Q72KP3_THET2	Q72kp3 thermus the
29	36	94.7	249	2	O22705_ARATH	O22705 arabidopsis
30	36	94.7	254	2	Q94AP2_ARATH	Q94ap2 arabidopsis
31	36	94.7	264	2	Q4WAY2_ASPFU	Q4way2 aspergillus
32	36	94.7	278	2	Q9L955_ZOORA	Q9l955 zoogloea ra
33	36	94.7	286	2	Q9M5P1_ORYSA	Q9m5p1 oryza sativ
34	36	94.7	286	2	Q7F613_ORYSA	Q7f613 oryza sativ
35	36	94.7	297	2	Q7WNM6_BORBR	Q7wnm6 bordetella
36	36	94.7	445	2	Q87JP5_VIBPA	Q87jp5 vibrio para
37	36	94.7	766	2	Q5B793_EMENI	Q5b793 aspergillus
38	36	94.7	766	2	Q96V54_EMENI	Q96v54 emericella
39	36	94.7	2384	2	Q4X1E3_ASPFU	Q4x1e3 aspergillus
40	35	92.1	108	2	Q5NWJ5_AZOSE	Q5nwj5 azoarcus sp
41	35	92.1	210	2	Q8PP64_XANAC	Q8pp64 xanthomonas
42	35	92.1	226	2	Q4TAI8_TETNG	Q4tai8 tetraodon n
43	35	92.1	254	2	Q5VPT2_ORYSA	Q5vpt2 oryza sativ
44	35	92.1	262	2	Q9Y991_AERPE	Q9y991 aeropyrum p
45	35	92.1	285	2	Q98IU4_RHILO	Q98iu4 rhizobium l

ALIGNMENTS

RESULT 1
Q5BB77_EMENI
ID Q5BB77_EMENI PRELIMINARY; PRT; 120 AA.
AC Q5BB77;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Predicted protein.
GN ORFNames=AN2203.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD01000035; EAA63860.1; -; Genomic_DNA.
SQ SEQUENCE 120 AA; 14016 MW; BF5160CDCB8924B3 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRLPPLP 7
|||||||
Db 51 SRLPPLP 57

```
RESULT 2
CREG1_HUMAN
ID   CREG1_HUMAN    STANDARD;      PRT;   220 AA.
AC   O75629; Q8N9A3;
DT   10-MAY-2005 (Rel. 47, Created)
DT   10-MAY-2005 (Rel. 47, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   CREG1 protein precursor (Cellular repressor of E1A-stimulated genes
DE   1).
GN   Name=CREG1; Synonyms=CREG; ORFNames=UNQ727/PRO1409;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA], AND FUNCTION.
RX   MEDLINE=98378515; PubMed=9710587;
RA   Veal E., Eisenstein M., Tseng Z.H., Gill G.;
RT   "A cellular repressor of E1A-stimulated genes that inhibits activation
RT   by E2F.";
RL   Mol. Cell. Biol. 18:5032-5041(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX   MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA   Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA   Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA   Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA   Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA   Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA   Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA   Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,
RA   Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA   Wood W.I., Godowski P.J., Gray A.M.;
RT   "The secreted protein discovery initiative (SPDI), a large-scale
RT   effort to identify novel human secreted and transmembrane proteins: a
RT   bioinformatics assessment.";
RL   Genome Res. 13:2265-2270(2003).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX   PubMed=14702039; DOI=10.1038/ng1285;
RA   Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA   Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA   Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA   Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA   Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA   Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA   Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA   Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA   Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA   Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA   Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA   Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA   Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA   Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA   Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA   Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA   Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA   Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA   Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA   Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA   Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA   Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA   Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA   Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA   Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA   Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT   "Complete sequencing and characterization of 21,243 full-length human
RT   cDNAs.";
RL   Nat. Genet. 36:40-45(2004).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RG   Human chromosome 1 international sequencing consortium;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Lymph, and Placenta;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   PROTEIN SEQUENCE OF 32-46.
RX   PubMed=15340161; DOI=10.1110/ps.04682504;
RA   Zhang Z., Henzel W.J.;
RT   "Signal peptide prediction based on analysis of experimentally
RT   verified cleavage sites.";
RL   Protein Sci. 13:2819-2824(2004).
RN   [7]
RP   SUBCELLULAR LOCATION, AND GLYCOSYLATION.
RX   MEDLINE=20273225; PubMed=10815803; DOI=10.1038/sj.onc.1203529;
RA   Veal E., Groisman R., Eisenstein M., Gill G.;
RT   "The secreted glycoprotein CREG enhances differentiation of NTERA-2
RT   human embryonal carcinoma cells.";
RL   Oncogene 19:2120-2128(2000).
RN   [8]
RP   INTERACTION WITH IGF2R, AND FUNCTION.
RX   MEDLINE=22815138; PubMed=12934103; DOI=10.1038/sj.onc.1206670;
RA   Di Bacco A., Gill G.;
RT   "The secreted glycoprotein CREG inhibits cell growth dependent on the
RT   mannose-6-phosphate/insulin-like growth factor II receptor.";
RL   Oncogene 22:5436-5445(2003).
CC   -!- FUNCTION: May contribute to the transcriptional control of cell
CC       growth and differentiation. Antagonizes transcriptional activation
CC       and cellular transformation by the adenovirus E1A protein. The
CC       transcriptional control activity of cell growth requires
CC       interaction with IGF2R.
CC   -!- SUBUNIT: Interacts with IGF2R; the interaction is dependent on
CC       glycosylation.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- PTM: N-glycosylated.
CC   -!- SIMILARITY: Belongs to the CREG family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AF084523; AAC34861.1; -; mRNA.
DR   EMBL; AY359071; AAQ89430.1; -; mRNA.
DR   EMBL; AK095456; BAC04550.1; -; mRNA.
DR   EMBL; AL031733; CAB42866.1; -; Genomic_DNA.
DR   EMBL; BC006786; AAH06786.1; -; mRNA.
DR   EMBL; BC008628; AAH08628.1; -; mRNA.
DR   Ensembl; ENSG00000143162; Homo sapiens.
```

```
DR   HGNC; HGNC:2351; CREG1.
DR   GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
DR   GO; GO:0003714; F:transcription corepressor activity; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0007275; P:development; TAS.
DR   GO; GO:0006357; P:regulation of transcription from RNA polyme. . .; TAS.
KW   Direct protein sequencing; Glycoprotein; Growth regulation; Signal.
FT   SIGNAL        1     31
FT   CHAIN        32    220       CREG1 protein.
FT   CARBOHYD    160    160       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    193    193       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    216    216       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT     52     59       Missing (in Ref. 4; BAC04550).
SQ   SEQUENCE   220 AA;  24075 MW;  0DB95A1E4149CD7C CRC64;

  Query Match             100.0%;  Score 38;  DB 1;  Length 220;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         44 SRLPPLP 50


RESULT 3
Q57XH6_9TRYP
ID   Q57XH6_9TRYP PRELIMINARY;      PRT;   318 AA.
AC   Q57XH6;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   RNA-binding protein, putative.
GN   ORFNames=Tb927.8.990;
OS   Trypanosoma brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5691;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=GUTat10.1;
RA   Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA   Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA   Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA   Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA   Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC159417; AAX69693.1; -; Genomic_DNA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0000398; P:nuclear mRNA splicing, via spliceosome; IEA.
DR   InterPro; IPR012677; a_b_plait_nuc_bd.
DR   InterPro; IPR000504; RNPl_RNA_bd.
DR   Pfam; PF00076; RRM_1; 1.
DR   SMART; SM00360; RRM; 1.
DR   PROSITE; PS50102; RRM; 1.
SQ   SEQUENCE   318 AA;  33963 MW;  B9368CC958B9E723 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 318;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db        166 SRLPPLP 172


RESULT 4
Q61ER4_CAEBR
ID   Q61ER4_CAEBR PRELIMINARY;      PRT;   358 AA.
AC   Q61ER4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein CBG11970 (Fragment).
GN   Name=CBG11970;
OS   Caenorhabditis briggsae.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6238;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RG   The C.briggsae Sequencing Consortium;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAC01000059; CAE66633.1; -; Genomic_DNA.
DR   GO; GO:0008168; F:methyltransferase activity; IEA.
DR   GO; GO:0000179; F:rRNA (adenine-N6,N6-)-dimethyltransferase a. . .; IEA.
DR   GO; GO:0008649; F:rRNA methyltransferase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0000154; P:rRNA modification; IEA.
DR   GO; GO:0006364; P:rRNA processing; IEA.
DR   InterPro; IPR001737; RRNA_meth_trans.
DR   Pfam; PF00398; RrnaAD; 1.
DR   SMART; SM00650; rADc; 1.
DR   PROSITE; PS01131; RRNA_A_DIMETH; 1.
KW   Hypothetical protein.
FT   NON_TER       1      1
FT   NON_TER     358    358
SQ   SEQUENCE   358 AA;  41285 MW;  A2A981D5F50800EC CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db          1 SRLPPLP 7


RESULT 5
UL29_HCMVA
ID   UL29_HCMVA     STANDARD;      PRT;   360 AA.
AC   P16764; Q7M6Q4;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Hypothetical protein UL29.
GN   Name=UL29;
OS   Human cytomegalovirus (strain AD169) (HHV-5) (Human herpesvirus 5).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
OX   NCBI_TaxID=10360;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX   MEDLINE=90269039; PubMed=2161319;
RA   Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA   Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
RA   Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT   "Analysis of the protein-coding content of the sequence of human
RT   cytomegalovirus strain AD169.";
RL   Curr. Top. Microbiol. Immunol. 154:125-169(1990).
RN   [2]
RP   GENOME REANNOTATION.
RX   MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;
RA   Davison A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
RA   Alcendor D.J., McGeoch D.J., Hayward G.S.;
RT   "The human cytomegalovirus genome revisited: comparison with the
RT   chimpanzee cytomegalovirus genome.";
RL   J. Gen. Virol. 84:17-28(2003).
RN   [3]
RP   ERRATUM.
RA   Davison A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
RA   Alcendor D.J., McGeoch D.J., Hayward G.S.;
RL   J. Gen. Virol. 84:1053-1053(2003).
CC   -!- SIMILARITY: Belongs to the herpesviruses US22 family.
CC   ----------------------------------------------------------------------
```

```
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   -----------------------------------------------------------------------
DR   EMBL; X17403; CAA35428.1; -; Genomic_DNA.
DR   EMBL; BK000394; DAA00133.1; -; Genomic_DNA.
DR   PIR; S09792; S09792.
DR   InterPro; IPR003360; US22.
DR   Pfam; PF02393; US22; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   360 AA;  40778 MW;  F989352FBD160004 CRC64;

  Query Match             100.0%;  Score 38;  DB 1;  Length 360;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         22 SRLPPLP 28


RESULT 6
Q6XNK8_HCMV
ID   Q6XNK8_HCMV PRELIMINARY;      PRT;   360 AA.
AC   Q6XNK8;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   UL29.
GN   Name=UL29; ORFNames=HHV5gp035;
OS   Human cytomegalovirus (HHV-5) (Human herpesvirus 5).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Cytomegalovirus.
OX   NCBI_TaxID=10359;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Towne;
RX   MEDLINE=22919658; PubMed=14557635;
RX   DOI=10.1128/JVI.77.21.11499-11506.2003;
RA   Komazin G., Ptak R.G., Emmer B.T., Townsend L.B., Drach J.C.;
RT   "Resistance of human cytomegalovirus to the benzimidazole L-
RT   ribonucleoside maribavir maps to UL27.";
RL   J. Virol. 77:11499-11506(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Toledo;
RA   Brondke H., Schmitz B., Shenk T., Doerfler W.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY223529; AAO73458.1; -; Genomic_DNA.
DR   EMBL; AY486471; AAS48939.1; -; Genomic_DNA.
DR   InterPro; IPR003360; US22.
DR   Pfam; PF02393; US22; 1.
SQ   SEQUENCE   360 AA;  40777 MW;  F989352FBD160004 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 360;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         22 SRLPPLP 28


RESULT 7
P91424_CAEEL
ID   P91424_CAEEL PRELIMINARY;      PRT;   367 AA.
AC   P91424;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=T03F1.7;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; U88169; AAB42235.2; -; Genomic_DNA.
DR   PIR; T29195; T29195.
DR   Ensembl; T03F1.7; Caenorhabditis elegans.
DR   WormBase; WBGene00020189; T03F1.7.
DR   WormPep; T03F1.7; CE30685.
DR   GO; GO:0000179; F:rRNA (adenine-N6,N6-)-dimethyltransferase a. . .; IEA.
DR   GO; GO:0008649; F:rRNA methyltransferase activity; IEA.
DR   GO; GO:0000154; P:rRNA modification; IEA.
DR   InterPro; IPR001737; RRNA_meth_trans.
DR   Pfam; PF00398; RrnaAD; 1.
DR   SMART; SM00650; rADc; 1.
DR   PROSITE; PS01131; RRNA_A_DIMETH; 1.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   367 AA;  41893 MW;  FD2419FC6548F1BC CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 367;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db          5 SRLPPLP 11


RESULT 8
Q9H8W6_HUMAN
ID   Q9H8W6_HUMAN PRELIMINARY;      PRT;   373 AA.
AC   Q9H8W6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein FLJ13171.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=14702039; DOI=10.1038/ng1285;
RA   Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA   Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA   Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA   Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA   Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA   Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA   Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA   Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA   Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA   Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA   Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA   Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA   Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA   Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA   Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA   Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA   Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA   Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA   Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA   Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
```

```
RA   Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA   Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA   Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA   Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA   Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA   Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT   "Complete sequencing and characterization of 21,243 full-length human
RT   cDNAs.";
RL   Nat. Genet. 36:40-45(2004).
DR   EMBL; AK023233; BAB14483.1; -; mRNA.
DR   InterPro; IPR004018; RPEL_repeat.
DR   Pfam; PF02755; RPEL; 2.
DR   SMART; SM00707; RPEL; 2.
SQ   SEQUENCE   373 AA;  43285 MW;  ECF59917ABADE459 CRC64;

  Query Match           100.0%;  Score 38;  DB 2;  Length 373;
  Best Local Similarity  100.0%;  Pred. No. 2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db         81 SRLPPLP 87


RESULT 9
Q5RAU1_PONPY
ID   Q5RAU1_PONPY PRELIMINARY;      PRT;   420 AA.
AC   Q5RAU1;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp469H1423.
GN   Name=DKFZp469H1423;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RG   The German cDNA Consortium;
RA   Ansorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR858921; CAH91119.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   420 AA;  44873 MW;  8BB463E5332A0A2A CRC64;

  Query Match           100.0%;  Score 38;  DB 2;  Length 420;
  Best Local Similarity  100.0%;  Pred. No. 2.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db        340 SRLPPLP 346


RESULT 10
O14498_HUMAN
ID   O14498_HUMAN PRELIMINARY;      PRT;   428 AA.
AC   O14498;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   ISLR precursor.
GN   Name=ISLR; ORFNames=UNQ189;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Caucasian; TISSUE=Retina;
RX   MEDLINE=97468140; PubMed=9325048; DOI=10.1006/geno.1997.4889;
RA   Nagasawa A., Kubota R., Imamura Y., Nagamine K., Wang  Y., Asakawa S.,
RA   Kudoh J., Minoshima S., Mashima Y., Oguchi Y., Shimizu N.;
RT   "Cloning of the cDNA for a new member of the immunoglobulin
RT   superfamily (ISLR) containing leucine-rich repeat (LRR).";
RL   Genomics 44:273-279(1997).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA   Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA   Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA   Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA   Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA   Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA   Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA   Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,
RA   Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA   Wood W.I., Godowski P.J., Gray A.M.;
RT   "The secreted protein discovery initiative (SPDI), a large-scale
RT   effort to identify novel human secreted and transmembrane proteins: a
RT   bioinformatics assessment.";
RL   Genome Res. 13:2265-2270(2003).
DR   EMBL; AB003184; BAA22848.1; -; mRNA.
DR   EMBL; AY358871; AAQ89230.1; -; mRNA.
DR   HSSP; P07359; 1M0Z.
DR   Ensembl; ENSG00000129009; Homo sapiens.
DR   HGNC; HGNC:6133; ISLR.
DR   GO; GO:0005515; F:protein binding; TAS.
DR   GO; GO:0007155; P:cell adhesion; TAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00560; LRR_1; 5.
DR   PRINTS; PR00019; LEURICHRPT.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin domain; Leucine-rich repeat; Repeat; Signal.
FT   SIGNAL        1     18       Potential.
SQ   SEQUENCE   428 AA;  45997 MW;  3163F89D59F6F3A4 CRC64;

  Query Match           100.0%;  Score 38;  DB 2;  Length 428;
  Best Local Similarity  100.0%;  Pred. No. 2.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db        222 SRLPPLP 228


RESULT 11
Q5NVQ6_PONPY
ID   Q5NVQ6_PONPY PRELIMINARY;      PRT;   428 AA.
AC   Q5NVQ6;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp459M1420.
GN   Name=DKFZp459M1420;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Cortex;
RG   The German cDNA Consortium;
RA   Bloecker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA   Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; CR925954; CAI29607.1; -; mRNA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR001611; LRR.
DR   InterPro; IPR000483; LRR_Cterm.
DR   InterPro; IPR003591; LRR_typ.
DR   Pfam; PF00560; LRR_1; 5.
DR   PRINTS; PR00019; LEURICHRPT.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00369; LRR_TYP; 5.
DR   SMART; SM00082; LRRCT; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Hypothetical protein; Immunoglobulin domain; Leucine-rich repeat;
KW   Repeat.
SQ   SEQUENCE   428 AA;  45824 MW;  A6A753F9E5363654 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 428;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db        222 SRLPPLP 228


RESULT 12
Q4R9C3_MACFA
ID   Q4R9C3_MACFA PRELIMINARY;      PRT;   517 AA.
AC   Q4R9C3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Testis cDNA clone: QtsA-10310, similar to human hypothetical protein
DE   FLJ13171 (FLJ13171),.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   International consortium for macaque cDNA sequencing, analysis;
RT   "DNA sequences of macaque genes expressed in brain or testis and its
RT   evolutionary implications.";
RL   Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA   Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT   "Substitution rate and structural divergence of 5'UTR evolution:
RT   Comparative analysis between human and cynomolgus monkey cDNAs.";
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB168173; BAE00298.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   517 AA;  57671 MW;  8C4B9A5085D9A784 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db        216 SRLPPLP 222


RESULT 13
Q20619_CAEEL
ID   Q20619_CAEEL PRELIMINARY;      PRT;   600 AA.
AC   Q20619;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein F49E12.6.
GN   ORFNames=F49E12.6;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR   EMBL; Z66520; CAA91391.2; -; Genomic_DNA.
DR   PIR; T22451; T22451.
DR   HSSP; Q16254; 1CF7.
DR   Ensembl; F49E12.6; Caenorhabditis elegans.
DR   WormBase; WBGene00009899; F49E12.6.
DR   WormPep; F49E12.6; CE37018.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0005667; C:transcription factor complex; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0000074; P:regulation of cell cycle; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   InterPro; IPR003316; E2F_TDP.
DR   InterPro; IPR011991; Wing_hlx_DNA_bd.
DR   Pfam; PF02319; E2F_TDP; 2.
KW   Complete proteome; DNA-binding; Hypothetical protein; Nuclear protein;
KW   Transcription; Transcription regulation.
SQ   SEQUENCE   600 AA;  66938 MW;  AEE4D04BE552A76F CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 600;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db        478 SRLPPLP 484


RESULT 14
Q6NUN6_HUMAN
ID   Q6NUN6_HUMAN PRELIMINARY;      PRT;   630 AA.
AC   Q6NUN6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   PHACTR4 protein.
GN   Name=PHACTR4;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Testis;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
```

```
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Testis;
RG   NIH MGC Project;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068508; AAH68508.1; -; mRNA.
DR   InterPro; IPR004018; RPEL_repeat.
DR   Pfam; PF02755; RPEL; 1.
DR   SMART; SM00707; RPEL; 2.
DR   PROSITE; PS51073; RPEL; 2.
SQ   SEQUENCE   630 AA;  69243 MW;  5FA0C75B4535C010 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 630;
  Best Local Similarity   100.0%;  Pred. No. 3.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db        394 SRLPPLP 400


RESULT 15
Q68DD4_HUMAN
ID   Q68DD4_HUMAN PRELIMINARY;      PRT;   654 AA.
AC   Q68DD4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein DKFZp686L07205.
GN   Name=DKFZp686L07205;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Fetal brain;
RG   The German cDNA Consortium;
RA   Bloecker H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA   Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR749449; CAH18286.1; -; mRNA.
DR   InterPro; IPR004018; RPEL_repeat.
DR   Pfam; PF02755; RPEL; 2.
DR   SMART; SM00707; RPEL; 3.
KW   Hypothetical protein.
SQ   SEQUENCE   654 AA;  72716 MW;  D2F7BE03A7D9F4E2 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SRLPPLP 7
              |||||||
Db        362 SRLPPLP 368


Search completed: April  4, 2006, 13:15:11
Job time : 7.35079 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 3.47251 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-293
Perfect score: 38
Sequence: 1 RALPSPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	7	3	AAB17237	Aab17237 SH3 antag
2	38	100.0	7	5	ABB73230	Abb73230 Src homol
3	38	100.0	7	7	ADJ73384	Adj73384 SH3 antag
4	38	100.0	7	8	ADJ53018	Adj53018 CH1 delet
5	38	100.0	7	8	ADJ51979	Adj51979 CH1 delet
6	38	100.0	13	2	AAW11120	Aaw11120 Src SH3 d
7	38	100.0	26	2	AAW16931	Aaw16931 Random re
8	38	100.0	26	2	AAW25494	Aaw25494 Random pe
9	38	100.0	78	4	AAG73409	Aag73409 Human gen
10	38	100.0	78	5	ABG64232	Abg64232 Human alb
11	38	100.0	78	8	ADL77497	Adl77497 Albumin f
12	38	100.0	216	6	ABP71204	Abp71204 S. cinnam
13	35	92.1	51	4	AAU49716	Aau49716 Propionib
14	35	92.1	51	6	ABM46235	Abm46235 Propionib
15	35	92.1	75	4	ABB15478	Abb15478 Human ner
16	35	92.1	84	3	AAG12401	Aag12401 Zea mays
17	35	92.1	109	5	ABP42067	Abp42067 Human ova
18	35	92.1	113	4	AAU53303	Aau53303 Propionib
19	35	92.1	113	6	ABM49822	Abm49822 Propionib
20	35	92.1	115	6	ADA54727	Ada54727 Human pro
21	35	92.1	123	3	AAB14305	Aab14305 Human sec
22	35	92.1	123	4	AAB85227	Aab85227 Human sec
23	35	92.1	204	7	ABO68038	Abo68038 Pseudomon
24	35	92.1	246	3	AAB58196	Aab58196 Lung canc
25	35	92.1	250	3	AAG50446	Aag50446 Arabidops
26	35	92.1	250	3	AAG14089	Aag14089 Arabidops
27	35	92.1	268	9	ABM94494	Abm94494 M. xanthu
28	35	92.1	278	4	AAU87155	Aau87155 Novel cen
29	35	92.1	278	8	ADI54470	Adi54470 Novel hum
30	35	92.1	298	7	ADM05717	Adm05717 Human pro
31	35	92.1	313	4	ABG09606	Abg09606 Novel hum
32	35	92.1	335	4	ABG18051	Abg18051 Novel hum
33	35	92.1	349	3	AAG14088	Aag14088 Arabidops
34	35	92.1	349	3	AAG50445	Aag50445 Arabidops
35	35	92.1	350	3	AAG50444	Aag50444 Arabidops
36	35	92.1	350	3	AAG14087	Aag14087 Arabidops
37	35	92.1	362	8	ADT56743	Adt56743 Plant pol
38	35	92.1	461	7	ABM86753	Abm86753 Rice abio
39	35	92.1	482	5	AAU98427	Aau98427 Cadium-re
40	35	92.1	483	3	AAG45923	Aag45923 Arabidops
41	35	92.1	483	3	AAG25784	Aag25784 Arabidops
42	35	92.1	485	3	AAG45922	Aag45922 Arabidops
43	35	92.1	485	3	AAG25783	Aag25783 Arabidops
44	35	92.1	485	4	AAB19934	Aab19934 Arabidops
45	35	92.1	485	6	ABP97714	Abp97714 Amino aci

ALIGNMENTS

RESULT 1
AAB17237
ID AAB17237 standard; peptide; 7 AA.
XX
AC AAB17237;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:293.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 299; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 1 RALPSPP 7

RESULT 2
ABB73230
ID ABB73230 standard; peptide; 7 AA.
XX
AC ABB73230;
XX
DT 05-APR-2002 (first entry)
XX
DE Src homology3 (SH3) antagonist peptide SEQ ID NO:293.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 55; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 1 RALPSPP 7

RESULT 3
ADJ73384
ID ADJ73384 standard; peptide; 7 AA.
XX
AC ADJ73384;
XX
DT 06-MAY-2004 (first entry)
XX
DE SH3 antagonist peptide sequence SeqID 839.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW SH3.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 839; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which

CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 1 RALPSPP 7

RESULT 4
ADJ53018
ID ADJ53018 standard; peptide; 7 AA.
XX
AC ADJ53018;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID839.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 839; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 1 RALPSPP 7

RESULT 5
ADJ51979
ID ADJ51979 standard; peptide; 7 AA.
XX
AC ADJ51979;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID839.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 839; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be

CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 1 RALPSPP 7

RESULT 6
AAW11120
ID AAW11120 standard; peptide; 13 AA.
XX
AC AAW11120;
XX
DT 25-JUN-1997 (first entry)
XX
DE Src SH3 domain-binding peptide used in signal transduction modulation.
XX
KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
OS Synthetic.
XX
PN WO9603649-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US009382.
XX
PR 22-JUL-1994; 94US-00278865.
PR 07-JUN-1995; 95US-00483555.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
XX
DR WPI; 1996-117151/12.
XX
PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.
XX
PS Claim 40; Page 83; 116pp; English.
XX
CC AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC binding peptides are useful in modulating signal transduction pathways at
CC the cellular level (especially protein tyrosine kinase-mediated),
CC modulating oncogenic protein activity, or providing compounds for the
CC development of drugs with the ability to modulate broad classes, as well
CC as specific classes, of proteins involved in signal transduction and also
CC for regulating the processing, trafficking or translation of RNA.
CC Conjugates of the peptides with detectable labels or imaging agents are
CC useful for imaging cells, tissues and organs in which Src or Src-related
CC proteins are expressed
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 38; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 4 RALPSPP 10

RESULT 7
AAW16931
ID AAW16931 standard; peptide; 26 AA.
XX
AC AAW16931;
XX
DT 27-JUN-1997 (first entry)
XX
DE Random recombinant SH3 domain binding peptide.
XX
KW Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW protein tyrosine kinase; signal transduction; RNA processing;
KW trafficking; translation.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /note= "X is undefined in the specification"
XX
PN WO9603649-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US009382.
XX
PR 22-JUL-1994; 94US-00278865.
PR 07-JUN-1995; 95US-00483555.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ;
XX
DR WPI; 1996-117151/12.
XX
PT Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT proteins - useful for e.g. modulating signal transduction pathways at the
PT cellular level, esp. protein tyrosine kinase-mediated.
XX
PS Disclosure; Fig 1; 116pp; English.
XX
CC AAW16924-W16948 are random recombinant peptides derived from one of three
CC peptide libraries, T9, T12 and R8C. The peptides are all SH3 domain-
CC binding peptides. SH3 binding peptides are useful in modulating signal
CC transduction pathways at the cellular level (especially protein tyrosine
CC kinase-mediated), modulating oncogenic protein activity, or providing
CC compounds for the development of drugs with the ability to modulate broad
CC classes, as well as specific classes, of proteins involved in signal
CC transduction and also for regulating the processing, trafficking or
CC translation of RNA. Conjugates of the peptides with detectable labels or
CC imaging agents are useful for imaging cells, tissues and organs in which
CC Src or Src-related proteins are expressed
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 38; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RALPSPP 7
              |||||||
Db         13 RALPSPP 19
```

RESULT 8
AAW25494
ID AAW25494 standard; peptide; 26 AA.
XX
AC AAW25494;
XX
DT 27-MAR-1998 (first entry)
XX
DE Random peptide recombinant clone T9.SRC3.7.
XX
KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS Synthetic.
OS Unidentified.
XX
FH Key   Location/Qualifiers
FT Misc-difference 1
FT   /note= "Any amino acid"
XX
PN WO9730074-A1.
XX
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-US002298.
XX
PR 16-FEB-1996; 96US-00602999.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;
XX
DR WPI; 1997-424972/39.
XX
PT Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.
XX
PS Disclosure; Fig 5; 131pp; English.
XX
CC The present sequence represents a random peptide recombinant isolated by
CC the method of the present invention. SH3 (Src homology region 3) binding
CC peptides are selected from: (a) peptides which bind the SH3 domain of
CC Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)
CC peptides which bind the SH3 domain of Abl; (d) peptides which bind the
CC SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;
CC (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind
CC the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3
CC domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain
CC of Grb2. The purified binding peptides can be used in the method to
CC identify inhibitors of their binding to their respective SH3 domains,
CC which could be used to modulate the pharmacological activity of proteins
CC or polypeptide containing the SH3 domain. The peptides can also be used
CC to activate Src or Src-related protein tyrosine kinases, to stimulate the
CC immune response by increasing the production of certain lymphokines, e.g.
CC tumour necrosis factor-alpha and interleukin-1, or to deliver a
CC conjugated molecule to certain cellular compartments containing Src or
CC Src related proteins
XX
SQ Sequence 26 AA;

```
  Query Match             100.0%;  Score 38;  DB 2;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 RALPSPP 7
              |||||||
Db         13 RALPSPP 19
```

RESULT 9
AAG73409
ID AAG73409 standard; protein; 78 AA.
XX
AC AAG73409;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human gene 14-encoded secreted protein HBXFZ38, SEQ ID NO:181.
XX
KW Human; secreted protein; proliferative disorder; cancer; chromosome 14;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder; tumour;
KW endocrine disorder; infection; wound healing; vulnerary; cell culture;
KW chemotaxis; food additive; binding partner identification.
XX
OS Homo sapiens.
XX
PN WO200134628-A1.
XX
PD 17-MAY-2001.
XX
PF 08-NOV-2000; 2000WO-US030653.
XX
PR 12-NOV-1999; 99US-0164735P.
PR 27-JUL-2000; 2000US-0221193P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Komatsoulis GA, Birse CE, Ni J, Moore PA;
XX
DR WPI; 2001-329066/34.
DR N-PSDB; AAH32586.
XX
PT Nucleic acids encoding 35 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease
PT and diabetic retinopathy.
XX
PS Claim 11; Page 543-544; 604pp; English.
XX
CC AAH32522-AAH32627 represent cDNAs corresponding to 35 human secreted
CC protein genes, and AAG73346-AAG73448 represent the proteins they encode.
CC AAG73449-AAG73519 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinsons's disease), cognitive disorders, schizophrenia, asthma, skin
CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their

CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein of the invention
XX
SQ Sequence 78 AA;

Query Match 100.0%; Score 38; DB 4; Length 78;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 59 RALPSPP 65

RESULT 10
ABG64232
ID ABG64232 standard; protein; 78 AA.
XX
AC ABG64232;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human albumin fusion protein #907.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US011988.
XX
PR 12-APR-2000; 2000US-0229358P.
PR 25-APR-2000; 2000US-0199384P.
PR 21-DEC-2000; 2000US-0256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2002-010886/01.
XX
PT New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein.
XX
PS Claim 1; Page 1092; 2102pp; English.
XX
CC The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or disorder
CC that may be modulated by therapeutic protein X. The albumin extends the
CC shelf-life of protein X, and may increase its biological in vitro/in vivo
CC activity. The protein is useful for treating and diagnosing disorders
CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
CC disease, ulcerative colitis), immune disorders (e.g. acquired
CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
CC ABG63326-ABG65518 represent albumin fusion proteins of the invention
XX
SQ Sequence 78 AA;

Query Match 100.0%; Score 38; DB 5; Length 78;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 59 RALPSPP 65

RESULT 11
ADL77497
ID ADL77497 standard; protein; 78 AA.
XX
AC ADL77497;
XX
DT 20-MAY-2004 (first entry)
XX
DE Albumin fusion protein related therapeutic protein X, SEQ ID No 979.
XX
KW albumin fusion protein; cytostatic; antianaemic; antiarthritic;
KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;
KW antipsoriatic; antibacterial; osteopathic; dermatological; antigout;
KW immunomodulator; antiarrhythmic; cardiant; nootropic; antilipaemic;
KW nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
KW antidiabetic; anabolic; hypertensive; vulnerary; gene therapy; cancer;
KW reproductive system disorder; therapeutic protein.
XX
OS Unidentified.
XX
PN US2004010134-A1.
XX
PD 15-JAN-2004.
XX
PF 12-APR-2001; 2001US-00833245.
XX
PR 12-APR-2000; 2000US-0229358P.
PR 25-APR-2000; 2000US-0199384P.
PR 21-DEC-2000; 2000US-0256931P.
XX
PA (ROSE/) ROSEN C A.
PA (HASE/) HASELTINE W A.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2004-090519/09.
XX
PT New albumin fusion proteins, useful for diagnosing, treating, preventing
PT or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,
PT asthma, inflammatory bowel disease or Alzheimer's disease.
XX
PS Disclosure; SEQ ID NO 979; 279pp; English.
XX
CC The invention relates to a novel albumin fusion protein. The invention
CC further relates to: a composition comprising the albumin fusion protein
CC and a pharmaceutical carrier; a kit comprising the composition of the
CC albumin fusion protein formula; a method of treating a disease or
CC disorder in a patient comprising the step of administering the albumin
CC fusion protein; a method of treating a patient with a disease or disorder
CC that is modulated by Therapeutic protein: X, or its fragment or variant;
CC a method of extending the shelf life of Therapeutic protein: X, or its
CC fragment or variant; a nucleic acid molecule comprising a polynucleotide
CC sequence encoding the albumin fusion protein; a vector comprising the
CC nucleic acid molecule of the albumin fusion protein; and a host cell
CC comprising the nucleic acid molecule of the albumin fusion protein. The
CC albumin fusion protein and its compositions have the following
CC activities: cytostatic, antianaemic, antiarthritic, antiasthmatic, anti-
CC HIV, immunosuppressive, antiinflammatory, antipsoriatic, antibacterial,
CC osteopathic, dermatological, antigout, immunomodulator, antiarrhythmic,

CC cardiant, nootropic, antilipaemic, nephrotropic, uropathic,
CC neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic,
CC hypertensive, and vulnerary. The albumin fusion protein nucleic acid may
CC be used in gene therapy to treat disorders. The albumin fusion protein is
CC useful for diagnosing, treating, preventing or ameliorating diseases or
CC disorders comprising indication: Y. The diseases or disorders include:
CC cancer (e.g. leukaemia, colon, bone, breast, liver or lung cancer),
CC immune or haematopoietic diseases (e.g. anaemia, Hodgkin's disease, acute
CC lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS,
CC autoimmune disease, inflammatory bowel disease, psoriasis or Lyme
CC disease), reproductive system disorders (e.g. prostatitis, inguinal
CC hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-
CC leydig tumours), musculoskeletal diseases (e.g. giant cell tumours,
CC Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy
CC or cachexia), cardiovascular disease (e.g. rhabdomyomas, heart disease,
CC arrhythmia, cardiac arrest, heat valve disease, hypernatraemia or
CC hyponatraemia), mixed foetal diseases (e.g. foetal alcohol syndrome,
CC Down's syndrome, Patau syndrome, Turner's syndrome, Apert syndrome or Tay
CC -Sachs disease), excretory diseases (e.g. urinary incontinence, urinary
CC tract infections or renal disorders), neural or sensory disease (e.g.
CC Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,
CC cerebellar ataxia, attention deficit disorder, autism or obsessive
CC compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or
CC occupational lung disease), endocrine diseases (e.g. diabetes, Addison's
CC disease or glomerulonephritis), digestive diseases (e.g. portal
CC hypertension, irritable bowel disease, gastric atrophy or pancreatitis)
CC or connective tissue or epithelial diseases (e.g. Crohn's disease,
CC scleroderma, wound healing or epidermolysis bullosa). This sequence
CC represents a therapeutic protein X relating to the albumin fusion protein
CC of the invention. The sequence listing data for this specification was
CC dowloaded from the USPTO website.
XX
SQ Sequence 78 AA;

Query Match 100.0%; Score 38; DB 8; Length 78;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 59 RALPSPP 65

RESULT 12
ABP71204
ID ABP71204 standard; protein; 216 AA.
XX
AC ABP71204;
XX
DT 14-APR-2003 (first entry)
XX
DE S. cinnamoneus cinnamycin cinR protein.
XX
KW Cinnamycin; bacterium; cinA; cinM; cinX; cinT; cinH; cinY; antibiotic;
KW food additive; antibacterial.
XX
OS Streptomyces cinnamoneus.
XX
PN WO200288367-A1.
XX
PD 07-NOV-2002.
XX
PF 29-APR-2002; 2002WO-GB001983.
XX
PR 27-APR-2001; 2001GB-00010432.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
PA (WIDD/) WIDDICK D A.
XX
PI Bibb MJ;
XX
DR WPI; 2003-111893/10.
DR N-PSDB; ABZ58812.
XX
PT New expression cassettes or genes isolated from Streptomyces cinnamoneus,
PT useful for producing a library of lantibiotic-producing host cells or
PT lantibiotics, which are useful as food additives and antibacterial
PT agents.
XX
PS Claim 13; Fig 20; 111pp; English.
XX
CC The invention relates to expression cassettes or sets of nucleic acids
CC comprising various open reading frames selected from (a) a cinA open
CC reading frame (orf), a cinM orf, and optionally a cinX orf; or (b) a cinA
CC orf, a cinM orf, a cinT orf, a cinH orf, a cinY orf, and optionally a
CC cinX orf. The expression cassettes, set of nucleic acids, (set of)
CC vectors, or methods are useful for producing a library of lantibiotic-
CC producing host cells or a library of lantibiotics. These are particularly
CC useful for producing lantibiotic cinnamycin or its modified versions. The
CC lantibiotics are useful as antibiotics having efficacy and utility as
CC food additives and antibacterial agents against Propionibacterium acnes
CC and problematic pathogens, e.g. methicillin-resistant Staphylococcus
CC aureus (which has or is developing resistance to many commonly used
CC antibiotics), or Streptococcus pneumoniae. Sequences ABP71191-71211
CC represent the various cinorf proteins encoded by the cinnamycin cluster
CC from S. cinnamoneus 40005 as present on the plasmid pDWFT9
XX
SQ Sequence 216 AA;

Query Match 100.0%; Score 38; DB 6; Length 216;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 70 RALPSPP 76

RESULT 13
AAU49716
ID AAU49716 standard; protein; 51 AA.
XX
AC AAU49716;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #10612.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59545.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

```
XX
PS   Example 1; SEQ ID NO 10911; 1069pp; English.
XX
CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC   polypeptides. The proteins and their associated DNA sequences are used in
CC   the treatment, prevention and diagnosis of medical conditions caused by
CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC   pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC   P. acnes is also involved in infections of bone, joints and the central
CC   nervous system, however it is particularly involved in the inflammatory
CC   lesions associated with acne vulgaris. A method for detecting the
CC   presence or absence of P. acnes in a patient comprises contacting a
CC   sample with a binding agent that binds to the proteins of the invention
CC   and determining the amount of bound protein in the sample. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   specific for P. acnes proteins. These antibodies can be used to
CC   downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 51 AA;

  Query Match             92.1%;  Score 35;  DB 4;  Length 51;
  Best Local Similarity   85.7%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||||:||
Db          9 RALPAPP 15


RESULT 14
ABM46235
ID   ABM46235 standard; protein; 51 AA.
XX
AC   ABM46235;
XX
DT   20-OCT-2003  (first entry)
XX
DE   Propionibacterium acnes predicted ORF-encoded polypeptide #10911.
XX
KW   Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW   immunostimulant; immune response; vaccine.
XX
OS   Propionibacterium acnes.
XX
PN   WO2003033515-A1.
XX
PD   24-APR-2003.
XX
PF   11-OCT-2002; 2002WO-US032727.
XX
PR   15-OCT-2001; 2001US-00978825.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Mitcham JL,  Skeiky YAW,  Persing DH,  Bhatia A,  Maisonneuve JL;
PI   Zhang Y,  Wang S,  Jen S,  Lodes MJ,  Benson DR,  Jones R,  Carter D;
PI   Barth B,  Vallieve-Douglass J;
XX
DR   WPI; 2003-381789/36.
DR   N-PSDB; ACF64474.
XX
PT   New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT   polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT   or for stimulating an immune response specific for a P. acnes protein.
XX
PS   Example 1; SEQ ID NO 10911; 1481pp; English.
XX
CC   The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC   encoding a Propionibacterium acnes protein. The invention also relates to
CC   polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC   immunogenic fragments of P. acnes polypeptides. The invention
CC   additionally encompasses expression vectors and host cells comprising a
CC   polynucleotide of the invention; antibodies against polypeptides of the
CC   invention; fusion proteins comprising a polypeptide of the invention; a
CC   method for stimulating an immune response specific for a P. acnes
CC   polypeptide and an isolated T cell population comprising T cells prepared
CC   via this method; a vaccine composition (comprising P. acnes polypeptides,
CC   polynucleotides, antibodies, fusion proteins, T cell populations, or
CC   antigen-presenting cells that express the polypeptide); a method and kit
CC   for detecting or determining the presence or absence of P. acnes in a
CC   patient; and a method for inhibiting the development of P. acnes in a
CC   patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC   proteins, T cell populations or antigen-presenting cells that express the
CC   polypeptides are useful for diagnosing, preventing or treating acne
CC   vulgaris, or for stimulating an immune response specific for a P. acnes
CC   protein. The polynucleotides can also be used as probes or primers for
CC   nucleic acid hybridisation. The vaccine composition is useful for the
CC   stimulation of an immune response against P. acnes, or for treating acne,
CC   and the kit is useful for performing a diagnostic assay. The present
CC   sequence represents a polypeptide predicted to be encoded by an ORF (open
CC   reading frame) contained within the P. acnes polynucleotides of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 51 AA;

  Query Match             92.1%;  Score 35;  DB 6;  Length 51;
  Best Local Similarity   85.7%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||||:||
Db          9 RALPAPP 15


RESULT 15
ABB15478
ID   ABB15478 standard; protein; 75 AA.
XX
AC   ABB15478;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human nervous system related polypeptide SEQ ID NO 4135.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001334.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
```

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-541565/60.
DR N-PSDB; ABA11804.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Claim 11; SEQ ID NO 4135; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.

CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 75 AA;

Query Match 92.1%; Score 35; DB 4; Length 75;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RALPSPP 7
              ||:||||
Db         35 RAVPSPP 41
```

Search completed: April 4, 2006, 13:07:42
Job time : 5.47251 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 1.14529 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-293
Perfect score: 38
Sequence: 1 RALPSPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	35	92.1	156	2	T37009	hypothetical prote
2	35	92.1	218	2	C64946	hypothetical prote
3	35	92.1	218	2	E85796	hypothetical prote
4	35	92.1	218	2	A99948	hypothetical prote
5	35	92.1	649	2	T01882	hypothetical prote
6	35	92.1	665	2	T04290	hypothetical prote
7	34	89.5	255	2	JG0179	superoxide dismuta
8	34	89.5	402	2	I46053	connexin44 - bovin
9	34	89.5	1222	2	G72614	probable reverse g
10	34	89.5	1571	2	T00062	hypothetical prote
11	34	89.5	3938	2	T42761	Bassoon protein -
12	34	89.5	3942	2	T42730	Bassoon protein -
13	33	86.8	135	2	T46448	hypothetical prote
14	33	86.8	169	2	T02081	ABA- and ripening-
15	33	86.8	196	2	T08808	hypothetical prote
16	33	86.8	204	2	AI0680	conserved hypothet
17	33	86.8	207	1	A64915	ycdY protein homol
18	33	86.8	207	2	F85764	probable oxidoredu
19	33	86.8	207	2	A90916	probable oxidoredu
20	33	86.8	236	2	H70708	probable ptrBb pro
21	33	86.8	314	2	AD0220	flagellar protein
22	33	86.8	469	2	B70201	hypothetical prote
23	33	86.8	496	2	T51058	hypothetical prote
24	33	86.8	529	2	H91012	hypothetical prote
25	33	86.8	529	2	C64987	probable oligopept
26	33	86.8	529	2	B85857	hypothetical prote
27	33	86.8	530	2	AD0155	probable ABC trans
28	33	86.8	546	2	T19680	hypothetical prote
29	33	86.8	642	2	T10861	phaseolin G-box bi
30	33	86.8	968	2	S46992	protein p130 - rat
31	33	86.8	999	2	I38547	novel cellular pro
32	33	86.8	3069	2	H70656	fatty-acid synthas
33	33	86.8	3076	2	A87058	fatty acid synthas
34	32	84.2	123	2	AH2707	conserved hypothet
35	32	84.2	141	2	G72641	hypothetical prote
36	32	84.2	223	2	A23036	nodulin-23 - soybe
37	32	84.2	224	2	S07315	nodulin - soybean
38	32	84.2	226	2	T36096	probable secreted
39	32	84.2	760	2	F86387	probable Pto kinas
40	32	84.2	1051	1	JW0051	serine/threonine-s
41	32	84.2	1133	2	T12529	hypothetical prote
42	32	84.2	1333	2	A37488	Ras guanine nucleo
43	32	84.2	1336	2	S25716	Ras guanine nucleo
44	32	84.2	1892	2	T18314	hypothetical prote
45	31	81.6	128	2	A75540	hypothetical prote

ALIGNMENTS

RESULT 1
T37009
hypothetical protein SCJ11.38c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37009
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21618
A;Accession: T37009
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-156 <OLI>
A;Cross-references: UNIPROT:Q9RI68; UNIPARC:UPI00000DB34D; EMBL:AL109949; PIDN:CAB5292
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ11.38c
C;Superfamily: Streptomyces coelicolor hypothetical protein SCJ11.38c

Query Match 92.1%; Score 35; DB 2; Length 156;
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
||:||||
Db 33 RAVPSPP 39

RESULT 2
C64946
hypothetical protein b1843 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: C64946
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C64946
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-218 <BLAT>
A;Cross-references: UNIPROT:P76280; UNIPARC:UPI000013BC61; GB:AE000278; GB:U00096; NID
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: Escherichia coli hypothetical protein b1843

Query Match 92.1%; Score 35; DB 2; Length 218;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7

```
              ||||:||
Db          55 RALPAPP 61


RESULT 3
E85796
hypothetical protein Z2893 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: E85796
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-218 <STO>
A;Cross-references: UNIPROT:Q8XCK7; UNIPARC:UPI0000165805; GB:AE005174; NID:g12515896; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z2893
C;Superfamily: Escherichia coli hypothetical protein b1843

  Query Match             92.1%;  Score 35;  DB 2;  Length 218;
  Best Local Similarity   85.7%;  Pred. No. 45;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||||:||
Db          55 RALPAPP 61


RESULT 4
A99948
hypothetical protein ECs2553 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A99948
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A99948
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-218 <HAY>
A;Cross-references: UNIPROT:Q8XCK7; UNIPARC:UPI00000D0475; GB:BA000007; PIDN:BAB35976.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2553
C;Superfamily: Escherichia coli hypothetical protein b1843

  Query Match             92.1%;  Score 35;  DB 2;  Length 218;
  Best Local Similarity   85.7%;  Pred. No. 45;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||||:||
Db          55 RALPAPP 61


RESULT 5
T01882
hypothetical protein F8M12.9 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01882
R;Madsen, C.; Graves, T.; Cotton, M.; Modde, T.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of A. thaliana F8M12.
A;Reference number: Z14450
A;Accession: T01882
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-649 <MAD>
A;Cross-references: UNIPROT:O81620; UNIPARC:UPI00000A3D11; EMBL:AF080118; NID:g3513725;
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 6/2; 201/1; 291/3; 372/1; 410/3; 490/3; 576/1
A;Note: F8M12.9

  Query Match             92.1%;  Score 35;  DB 2;  Length 649;
  Best Local Similarity   85.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||||:||
Db        223 RALPAPP 229


RESULT 6
T04290
hypothetical protein F25I24.160 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04290
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15261
A;Accession: T04290
A;Molecule type: DNA
A;Residues: 1-665 <BEV>
A;Cross-references: UNIPROT:Q9SN59; UNIPARC:UPI00000AA1C9; EMBL:AL049525
A;Experimental source: cultivar Columbia; BAC clone F25I24
C;Genetics:
A;Map position: 4
A;Introns: 22/2; 218/2; 307/3; 388/1; 426/3; 506/3; 592/1
A;Note: F25I24.160

  Query Match             92.1%;  Score 35;  DB 2;  Length 665;
  Best Local Similarity   85.7%;  Pred. No. 1.4e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||||:||
Db        239 RALPAPP 245


RESULT 7
JG0179
superoxide dismutase (EC 1.15.1.1) (Fe) - rice
C;Species: Oryza sativa (rice)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C;Accession: JG0179
R;Kaminaka, H.; Morita, S.; Tokumoto, M.; Yokoyama, H.; Masumura, T.; Tanaka, K.
Biosci. Biotechnol. Biochem. 63, 302-308, 1999
A;Title: Molecular cloning and characterization of a cDNA for an iron-superoxide dismut
A;Reference number: JG0179; MUID:99208990; PMID:10192910
A;Accession: JG0179
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-255 <KAM>
A;Cross-references: UNIPROT:Q9ZWM8; UNIPARC:UPI00000A6CFB; DDBJ:AB014056; NID:g4164148;
C;Superfamily: superoxide dismutase (Mn)
C;Keywords: iron; metalloprotein; oxidoreductase
F;67,119,203,207/Binding site: iron (His, His, Asp, His) #status predicted

  Query Match             89.5%;  Score 34;  DB 2;  Length 255;
  Best Local Similarity   85.7%;  Pred. No. 79;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
```

```
              | |||||
Db          9 RVLPSPP 15


RESULT 8
I46053
connexin44 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 13-Aug-1999
C;Accession: I46053
R;Gupta, V.K.; Berthoud, V.M.; Atal, N.; Jarillo, J.A.; Barrio, L.C.; Beyer, E.C.
Invest. Ophthalmol. Vis. Sci. 35, 3747-3758, 1994
A;Title: Bovine connexin44, a lens gap junction protein: molecular cloning, immunologica
A;Reference number: I46053; MUID:94375220; PMID:8088962
A;Accession: I46053
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-402 <GUP>
A;Cross-references: UNIPARC:UPI0000177531; EMBL:U08213; NID:g469557; PIDN:AAA50954.1; PI
C;Superfamily: gap junction protein

  Query Match             89.5%;  Score 34;  DB 2;  Length 402;
  Best Local Similarity   85.7%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||| ||
Db        292 RALPGPP 298


RESULT 9
G72614
probable reverse gyrase APE1376 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G72614
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: G72614
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1222 <KAW>
A;Cross-references: UNIPROT:Q9YC75; UNIPARC:UPI000005DF12; DDBJ:AP000061; NID:g5104821;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1376

  Query Match             89.5%;  Score 34;  DB 2;  Length 1222;
  Best Local Similarity   85.7%;  Pred. No. 4.1e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              | |||||
Db         40 RCLPSPP 46


RESULT 10
T00062
hypothetical protein KIAA0434 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C;Accession: T00062
R;Ishikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka, A.;
submitted to the EMBL Data Library, October 1997
A;Description: Prediction of the coding sequences of unidentified human genes. VIII. The
A;Reference number: Z14082
A;Accession: T00062
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1571 <ISH>
A;Cross-references: UNIPROT:O43161; UNIPARC:UPI00001693CF; EMBL:AB007894; NID:g2662148
A;Experimental source: brain; clone HH2165
C;Genetics:
A;Note: KIAA0434

  Query Match             89.5%;  Score 34;  DB 2;  Length 1571;
  Best Local Similarity   85.7%;  Pred. No. 5.3e+02;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              | |||||
Db        540 RTLPSPP 546


RESULT 11
T42761
Bassoon protein - rat
N;Alternate names: brain-specific synapse-associated protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42761
R;Dieck, S.; Sanmarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; We
J. Cell Biol. 142, 499-509, 1998
A;Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localiz
A;Reference number: Z22249; MUID:98345363; PMID:9679147
A;Accession: T42761
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3938 <DIE>
A;Cross-references: UNIPROT:O88778; UNIPARC:UPI00000E7EDE; EMBL:Y16563; NID:g3413503;
A;Experimental source: strain Sprague Dawley; brain
C;Function:
A;Description: may be involved in cytomatrix organization at the site of neurotransmit
A;Note: component of the presynaptic cytoskeleton
C;Keywords: coiled coil; zinc finger

  Query Match             89.5%;  Score 34;  DB 2;  Length 3938;
  Best Local Similarity   85.7%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              | |||||
Db       2889 RTLPSPP 2895


RESULT 12
T42730
Bassoon protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42730
R;Dieck, S.; Sanmarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; We
J. Cell Biol. 142, 499-509, 1998
A;Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localiz
A;Reference number: Z22249; MUID:98345363; PMID:9679147
A;Accession: T42730
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-3942 <DIE>
A;Cross-references: UNIPROT:O88737; UNIPARC:UPI0000029B58; EMBL:Y17034; NID:g3413809;
A;Experimental source: strain 129 SVJ
C;Genetics:
A;Map position: 9F1
A;Introns: 72/2; 208/3; 505/3; 675/3; 2889/3; 3582/1; 3851/3; 3886/1; 3930/1
A;Note: bassoon
C;Function:
A;Description: may be involved in cytomatrix organization at the site of neurotransmit
A;Note: component of the presynaptic cytoskeleton
C;Keywords: coiled coil; zinc finger

  Query Match             89.5%;  Score 34;  DB 2;  Length 3942;
  Best Local Similarity   85.7%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 RALPSPP 7
              | |||||
Db       2904 RTLPSPP 2910
```

RESULT 13
T46448
hypothetical protein DKFZp434N1429.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46448
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23037
A;Accession: T46448
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-135 <AAA>
A;Cross-references: UNIPROT:Q9NTF5; UNIPARC:UPI000006D41B; EMBL:AL137301
A;Experimental source: adult testis; clone DKFZp434N1429
C;Genetics:
A;Note: DKFZp434N1429.1

```
  Query Match             86.8%;  Score 33;  DB 2;  Length 135;
  Best Local Similarity   85.7%;  Pred. No. 60;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||| ||
Db        100 RALPRPP 106
```

RESULT 14
T02081
ABA- and ripening-inducible-like protein - maize
C;Species: Zea mays (maize)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C;Accession: T02081
R;Arredondo-Peter, R.; Shearman, L.; Ji, L.; Klucas, R.V.
submitted to the EMBL Data Library, April 1994
A;Description: Nucleotide sequence of an ABA- and ripening-like cDNA isolated from corn
A;Reference number: Z14553
A;Accession: T02081
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-169 <ARR>
A;Cross-references: UNIPROT:Q41730; UNIPARC:UPI00000A2CC1; EMBL:U09276; NID:g551482; PID
A;Experimental source: strain Golden Bantam; mesophyll

```
  Query Match             86.8%;  Score 33;  DB 2;  Length 169;
  Best Local Similarity   85.7%;  Pred. No. 76;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              || ||||
Db         64 RAAPSPP 70
```

RESULT 15
T08808
hypothetical protein DKFZp586J1923.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08808
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16472
A;Accession: T08808
A;Molecule type: mRNA
A;Residues: 1-196 <ANS>
A;Cross-references: UNIPROT:Q9UKR3; UNIPARC:UPI000016AC51; EMBL:AL050220
A;Experimental source: adult uterus; clone DKFZp586J1923
C;Genetics:
A;Note: DKFZp586J1923.1
C;Superfamily: trypsin; trypsin homology

```
  Query Match             86.8%;  Score 33;  DB 2;  Length 196;
  Best Local Similarity   100.0%;  Pred. No. 89;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ALPSPP 7
              ||||||
Db          7 ALPSPP 12
```

Search completed: April 4, 2006, 13:17:25
Job time : 3.14529 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 5.35079 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-293
Perfect score: 38
Sequence: 1 RALPSPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	79	2	Q6Z3R7_ORYSA	Q6z3r7 oryza sativ
2	38	100.0	216	2	Q83VX9_STRCJ	Q83vx9 streptomyce
3	38	100.0	233	2	Q6ZIV3_ORYSA	Q6ziv3 oryza sativ
4	38	100.0	256	2	Q8H2Z5_ORYSA	Q8h2z5 oryza sativ
5	38	100.0	408	2	Q94LF5_ORYSA	Q94lf5 oryza sativ
6	38	100.0	543	2	Q5WMR3_ORYSA	Q5wmr3 oryza sativ
7	38	100.0	1245	2	Q55ZE2_CRYNE	Q55ze2 cryptococcu
8	38	100.0	1245	2	Q5KNQ7_CRYNE	Q5knq7 cryptococcu
9	38	100.0	3089	2	Q6XXM0_MYCSM	Q6xxm0 mycobacteri
10	36	94.7	74	2	Q67UA0_ORYSA	Q67ua0 oryza sativ
11	36	94.7	389	2	Q6Z8N1_ORYSA	Q6z8n1 oryza sativ
12	36	94.7	398	2	Q52GR5_MAGGR	Q52gr5 magnaporthe
13	35	92.1	122	2	Q5ENX5_9VIRU	Q5enx5 torque teno
14	35	92.1	148	2	Q4R982_MACFA	Q4r982 macaca fasc
15	35	92.1	150	2	Q9WQH1_9VIRU	Q9wqh1 torque teno
16	35	92.1	154	2	Q9IFV1_9VIRU	Q9ifv1 torque teno
17	35	92.1	156	2	Q9RI68_STRCO	Q9ri68 streptomyce
18	35	92.1	160	2	Q9WB02_9VIRU	Q9wb02 torque teno
19	35	92.1	173	2	Q7XQ86_ORYSA	Q7xq86 oryza sativ
20	35	92.1	218	1	YOBB_ECOLI	P76280 escherichia
21	35	92.1	218	2	Q7AD77_ECO57	Q7ad77 escherichia
22	35	92.1	218	2	Q8FGS4_ECOL6	Q8fgs4 escherichia
23	35	92.1	218	2	Q8XCK7_ECO57	Q8xck7 escherichia
24	35	92.1	244	2	Q504R6_HUMAN	Q504r6 homo sapien
25	35	92.1	250	2	Q682T5_ARATH	Q682t5 arabidopsis
26	35	92.1	257	2	Q5ENX7_9VIRU	Q5enx7 torque teno
27	35	92.1	266	2	Q5ENX6_9VIRU	Q5enx6 torque teno
28	35	92.1	268	1	NO20_MEDTR	P93329 medicago tr
29	35	92.1	277	2	Q6Z6Q2_ORYSA	Q6z6q2 oryza sativ
30	35	92.1	288	2	P91249_CAEEL	P91249 caenorhabdi
31	35	92.1	298	2	Q8N7Z6_HUMAN	Q8n7z6 homo sapien
32	35	92.1	335	2	Q5F1X7_ORYSA	Q5f1x7 oryza sativ
33	35	92.1	350	2	Q680C0_ARATH	Q680c0 arabidopsis
34	35	92.1	390	2	Q8KLK5_STRTO	Q8klk5 streptomyce
35	35	92.1	409	2	Q5Z8A4_ORYSA	Q5z8a4 oryza sativ
36	35	92.1	485	2	Q698X7_9BRAS	Q698x7 thlaspi cae
37	35	92.1	485	2	Q852U0_BRAJU	Q852u0 brassica ju
38	35	92.1	485	2	Q8LST0_9BRAS	Q8lst0 thlaspi cae
39	35	92.1	485	2	Q8LST1_9BRAS	Q8lst1 thlaspi jap
40	35	92.1	485	2	Q8W122_ARATH	Q8w122 arabidopsis
41	35	92.1	485	2	Q94IP0_BRAJU	Q94ip0 brassica ju
42	35	92.1	485	2	Q9M6R0_ARATH	Q9m6r0 arabidopsis
43	35	92.1	485	2	Q9S7Z3_ARATH	Q9s7z3 arabidopsis
44	35	92.1	485	2	Q9ZPM2_ARATH	Q9zpm2 arabidopsis
45	35	92.1	485	2	Q56VL5_ARAHA	Q56vl5 arabidopsis

ALIGNMENTS

RESULT 1
Q6Z3R7_ORYSA
ID Q6Z3R7_ORYSA PRELIMINARY; PRT; 79 AA.
AC Q6Z3R7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNBa0025J22.39.
GN Name=OSJNBa0025J22.39;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP005245; BAD10187.1; -; Genomic_DNA.
DR Gramene; Q6Z3R7; -.
KW Hypothetical protein.
SQ SEQUENCE 79 AA; 7710 MW; 21152EE2F5B44C7A CRC64;

Query Match 100.0%; Score 38; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 43 RALPSPP 49

RESULT 2
Q83VX9_STRCJ
ID Q83VX9_STRCJ PRELIMINARY; PRT; 216 AA.
AC Q83VX9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CinR protein.
GN Name=cinR;
OS Streptomyces cinnamoneus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=53446;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Type strain DSM 40005;
RX MEDLINE=22558541; PubMed=12642677; DOI=10.1073/pnas.0230516100;
RA Widdick D.A., Dodd H.M., Barraille P., White J., Stein T.H.,
RA Chater K.F., Gasson M.J., Bibb M.J.;
RT "Cloning and engineering of the cinnamycin biosynthetic cluster from
RT Streptomyces cinnamoneus cinnamoneus DSM40005.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:4316-4321(2003).

```
CC   -!- SIMILARITY: Contains 1 HTH luxR-type DNA-binding domain.
DR   EMBL; AJ536588; CAD60529.1; -; Genomic_DNA.
DR   HSSP; P10957; 1RNL.
DR   GO; GO:0005622; C:intracellular; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0000156; F:two-component response regulator activity; IEA.
DR   GO; GO:0007600; P:sensory perception; IEA.
DR   GO; GO:0006350; P:transcription; IEA.
DR   GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR   InterPro; IPR000792; HTH_LuxR.
DR   InterPro; IPR001789; Response_reg.
DR   InterPro; IPR011991; Wing_hlx_DNA_bd.
DR   Pfam; PF00196; GerE; 1.
DR   Pfam; PF00072; Response_reg; 1.
DR   PRINTS; PR00038; HTHLUXR.
DR   ProDom; PD000307; HTH_LuxR; 1.
DR   ProDom; PD000039; Response_reg; 1.
DR   SMART; SM00421; HTH_LUXR; 1.
DR   SMART; SM00448; REC; 1.
DR   PROSITE; PS50110; RESPONSE_REGULATORY; 1.
KW   DNA-binding; Sensory transduction; Transcription;
KW   Transcription regulation; Two-component regulatory system.
SQ   SEQUENCE   216 AA;  22938 MW;  5D1E66581B306AB9 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db         70 RALPSPP 76


RESULT 3
Q6ZIV3_ORYSA
ID   Q6ZIV3_ORYSA PRELIMINARY;      PRT;   233 AA.
AC   Q6ZIV3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein OJ1198_B10.22 (Hypothetical protein
DE   OJ1051_A08.10).
GN   Name=OJ1198_B10.22; Synonyms=OJ1051_A08.10;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC
RT   clone:OJ1198_B10.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC
RT   clone:OJ1051_A08.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP003947; BAC99436.1; -; Genomic_DNA.
DR   EMBL; AP003904; BAC99374.1; -; Genomic_DNA.
DR   Gramene; Q6ZIV3; -.
KW   Hypothetical protein.
SQ   SEQUENCE   233 AA;  24114 MW;  5702BBEA55F176D7 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db         17 RALPSPP 23


RESULT 4
Q8H2Z5_ORYSA
ID   Q8H2Z5_ORYSA PRELIMINARY;      PRT;   256 AA.
AC   Q8H2Z5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Root cap protein 1-like.
GN   Name=P0710F09.141;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
RT   clone:P0710F09.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005325; BAC21552.1; -; Genomic_DNA.
DR   Gramene; Q8H2Z5; -.
SQ   SEQUENCE   256 AA;  27126 MW;  4A717F5DB75042C1 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 256;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db        141 RALPSPP 147


RESULT 5
Q94LF5_ORYSA
ID   Q94LF5_ORYSA PRELIMINARY;      PRT;   408 AA.
AC   Q94LF5; Q84MW4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein (Expressed protein).
GN   ORFNames=OSJNBa0015K03.1, Os03g28130;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Buell C., Yuan Q., Ouyang S., Moffat K., Hill J., Gansberger K.,
RA   Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA   Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., VanAken S.,
RA   Utterback T., Feldblyum T., Quackenbush J., Salzberg S., White O.,
RA   Fraser C.;
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Buell R.;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Buell C., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.,
RA   Overton II L., Tsitrin T., Kim M., Bera J., Jin S., Fadrosh D.,
RA   Tallon L., Koo H., Zismann V., Hsiao J., Blunt S., Vanaken S.,
RA   Riedmuller S., Utterback T., Feldblyum T., Yang Q., Haas B., Suh B.,
RA   Peterson J., Quackenbush J., White O., Salzberg S., Fraser C.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC084295; AAK55468.1; -; Genomic_DNA.
DR   EMBL; AC091787; AAP12928.1; -; Genomic_DNA.
DR   Gramene; Q84MW4; -.
DR   Gramene; Q94LF5; -.
```

```
DR   GO; GO:0006512; P:ubiquitin cycle; IEA.
DR   InterPro; IPR001810; F-box.
DR   Pfam; PF00646; F-box; 1.
DR   PROSITE; PS50181; FBOX; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   408 AA;  42936 MW;  89E9AC2946B18794 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 408;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db        142 RALPSPP 148


RESULT 6
Q5WMR3_ORYSA
ID   Q5WMR3_ORYSA PRELIMINARY;      PRT;   543 AA.
AC   Q5WMR3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein OJ1123_C08.7.
GN   Name=OJ1123_C08.7;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA   Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA   Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA   Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA   Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA   Wu H.-P., Shaw J.-F.;
RT   "Oryza sativa BAC OJ1123_C08 genomic sequence.";
RL   Submitted (OCT-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC108875; AAV32132.1; -; Genomic_DNA.
DR   InterPro; IPR007719; Phytcheltn_synth.
DR   Pfam; PF05023; Phytochelatin; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   543 AA;  59362 MW;  C843FEEAA9F6C9EE CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 543;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db         73 RALPSPP 79


RESULT 7
Q55ZE2_CRYNE
ID   Q55ZE2_CRYNE PRELIMINARY;      PRT;  1245 AA.
AC   Q55ZE2;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=CNBA5470;
OS   Cryptococcus neoformans var. neoformans B-3501A.
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC   Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX   NCBI_TaxID=283643;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=B-3501A;
RA   Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA   Wickes B.L., Fu J., Davis R.W.;
RT   "Cryptococcus neoformans serotype D sequencing.";
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAEY01000003; EAL23203.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1245 AA;  133752 MW;  B43A4C45FDD8926F CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 1245;
  Best Local Similarity   100.0%;  Pred. No. 7.9e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db        341 RALPSPP 347


RESULT 8
Q5KNQ7_CRYNE
ID   Q5KNQ7_CRYNE PRELIMINARY;      PRT;  1245 AA.
AC   Q5KNQ7;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=CNA05670;
OS   Cryptococcus neoformans var. neoformans JEC21.
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC   Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX   NCBI_TaxID=214684;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=JEC21;
RA   Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA   Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA   Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA   D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA   Huang J.C., Janbon G., Jones S.J.M., Krzywinski M.I., Kwon-Chung J.K.,
RA   Lengeler K.B., Maiti R., Marra M.A., Marra R.E., Mathewson C.A.,
RA   Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L., Shvartsbeyn A.,
RA   Schein J.E., Shin H., Specht C.A., Suh B., Tenney A., Utterback T.,
RA   Wickes B.L., Wye N.H., Kronstad J., Lodge J.K., Heitman J.,
RA   Davis R.W., Fraser C.M., Hyman R.W.;
RT   "The genome and transcriptome of Cryptococcus neoformans, a
RT   basidiomycete fungal pathogen of humans.";
RL   Science 0:0-0(2005).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=JEC21;
RA   Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,
RA   Van Aken S., Fraser C.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=JEC21;
RX   PubMed=15653466; DOI=10.1126/science.1103773;
RA   Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA   Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA   Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA   D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA   Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA   Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA   Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
RA   Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA   Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA   Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA   Fraser C.M., Hyman R.W.;
RT   "The genome of the basidiomycetous yeast and human pathogen
RT   Cryptococcus neoformans.";
RL   Science 307:1321-1324(2005).
DR   EMBL; AE017341; AAW41070.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
```

```
SQ   SEQUENCE   1245 AA;  133753 MW;  B43A4C45FDD8926F CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 1245;
  Best Local Similarity   100.0%;  Pred. No. 7.9e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db        341 RALPSPP 347


RESULT 9
Q6XXM0_MYCSM
ID   Q6XXM0_MYCSM PRELIMINARY;      PRT;  3089 AA.
AC   Q6XXM0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Fatty acid synthetase I.
GN   Name=fas1;
OS   Mycobacterium smegmatis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1772;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Zimhony O., Vilcheze C., Jacobs W.R. Jr.;
RT   "Characterization of Mycobacterium smegmatis Expressing the
RT   Mycobacterium tuberculosis Fatty Acid Synthase I (fas1) Gene.";
RL   J. Bacteriol. 186:4051-4055(2004).
DR   EMBL; AY205337; AAO43178.1; -; Genomic_DNA.
DR   GO; GO:0005835; C:fatty-acid synthase complex; IEA.
DR   GO; GO:0004312; F:fatty-acid synthase activity; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR001227; Ac_transferase.
DR   InterPro; IPR003965; Fatty_acid_synth.
DR   InterPro; IPR000794; Ketoacyl_synth.
DR   InterPro; IPR002539; MaoC_dehydratas.
DR   Pfam; PF00698; Acyl_transf_1; 1.
DR   Pfam; PF00109; ketoacyl-synt; 1.
DR   Pfam; PF02801; Ketoacyl-synt_C; 1.
DR   Pfam; PF01575; MaoC_dehydratas; 1.
DR   PRINTS; PR01483; FASYNTHASE.
DR   PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
SQ   SEQUENCE   3089 AA;  329440 MW;  F61433A3A824A40A CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 3089;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+03;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db       2433 RALPSPP 2439


RESULT 10
Q67UA0_ORYSA
ID   Q67UA0_ORYSA PRELIMINARY;      PRT;    74 AA.
AC   Q67UA0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein P0025H07.9.
GN   Name=P0025H07.9;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC
RT   clone:P0025H07.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005655; BAD38271.1; -; Genomic_DNA.
DR   Gramene; Q67UA0; -.
KW   Hypothetical protein.
SQ   SEQUENCE   74 AA;  8159 MW;  6D5E080B2B429E06 CRC64;

  Query Match             94.7%;  Score 36;  DB 2;  Length 74;
  Best Local Similarity   85.7%;  Pred. No. 89;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||:||||
Db         55 RAMPSPP 61


RESULT 11
Q6Z8N1_ORYSA
ID   Q6Z8N1_ORYSA PRELIMINARY;      PRT;   389 AA.
AC   Q6Z8N1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein P0711H09.20.
GN   Name=P0711H09.20;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP004765; BAD10070.1; -; Genomic_DNA.
DR   Gramene; Q6Z8N1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   389 AA;  45149 MW;  F418722F6750F4DC CRC64;

  Query Match             94.7%;  Score 36;  DB 2;  Length 389;
  Best Local Similarity   85.7%;  Pred. No. 5.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||:||||
Db         14 RAIPSPP 20


RESULT 12
Q52GR5_MAGGR
ID   Q52GR5_MAGGR PRELIMINARY;      PRT;   398 AA.
AC   Q52GR5;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=MG01221.4;
OS   Magnaporthe grisea 70-15.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX   NCBI_TaxID=242507;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=70-15;
RA   Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA   Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA   Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA   Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA   Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA   Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
```

```
RA   Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA   David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA   Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA   Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA   Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA   Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA   Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA   Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,
RA   Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA   Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA   Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA   Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA   Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA   Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA   Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA   Mozes J., Mulrain L., Munson G., Naylor J., Newes C., Nguyen C.,
RA   Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA   Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA   O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piqani B.,
RA   Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA   Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA   Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA   Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA   Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA   Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA   Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA   Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA   Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA   Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA   Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA   Zimmer A., Zody M., Lander E.;
RT   "The genome sequence of Magnaporthe grisea.";
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=70-15;
RA   Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=70-15;
RA   Zhu H., Blackmon B.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACU01000013; EAA55570.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   398 AA;  43544 MW;  A67E5FF9249DF4C6 CRC64;

  Query Match             94.7%;  Score 36;  DB 2;  Length 398;
  Best Local Similarity   85.7%;  Pred. No. 5.2e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||:||||
Db        220 RAMPSPP 226


RESULT 13
Q5ENX5_9VIRU
ID   Q5ENX5_9VIRU PRELIMINARY;      PRT;   122 AA.
AC   Q5ENX5;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   ORF2.
OS   Torque teno virus.
OC   Viruses; ssDNA viruses; Anellovirus.
OX   NCBI_TaxID=68887;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=2h;
RX   PubMed=15831945; DOI=10.1099/vir.0.80794-0;
RA   Niel C., Diniz-Mendes L., Devalle S.;
RT   "Rolling-circle amplification of Torque teno virus (TTV) complete
RT   genomes from human and swine sera and identification of a novel swine
RT   TTV genogroup.";
RL   J. Gen. Virol. 86:1343-1347(2005).
DR   EMBL; AY823988; AAW79274.1; -; Genomic_DNA.
DR   InterPro; IPR004118; TT_ORF2.
DR   Pfam; PF02957; TT_ORF2; 1.
SQ   SEQUENCE   122 AA;  12789 MW;  3B2008325B8314A5 CRC64;

  Query Match             92.1%;  Score 35;  DB 2;  Length 122;
  Best Local Similarity   85.7%;  Pred. No. 2.2e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||||:||
Db         65 RALPAPP 71


RESULT 14
Q4R982_MACFA
ID   Q4R982_MACFA PRELIMINARY;      PRT;   148 AA.
AC   Q4R982;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Testis cDNA clone: QtsA-10553, similar to human checkpoint suppressor
DE   1 (CHES1),.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   International consortium for macaque cDNA sequencing, analysis;
RT   "DNA sequences of macaque genes expressed in brain or testis and its
RT   evolutionary implications.";
RL   Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA   Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT   "Substitution rate and structural divergence of 5'UTR evolution:
RT   Comparative analysis between human and cynomolgus monkey cDNAs.";
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB168214; BAE00339.1; -; mRNA.
SQ   SEQUENCE   148 AA;  15574 MW;  E1A60DC0DDAA73FB CRC64;

  Query Match             92.1%;  Score 35;  DB 2;  Length 148;
  Best Local Similarity   85.7%;  Pred. No. 2.7e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||:||||
Db         31 RAVPSPP 37


RESULT 15
Q9WQH1_9VIRU
ID   Q9WQH1_9VIRU PRELIMINARY;      PRT;   150 AA.
AC   Q9WQH1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Torque teno virus.
OC   Viruses; ssDNA viruses; Anellovirus.
OX   NCBI_TaxID=68887;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
```

```
RX   MEDLINE=99179035; PubMed=10077657; DOI=10.1073/pnas.96.6.3177;
RA   Mushahwar I.K., Erker J.C., Muerhoff A.S., Leary T.P., Simons J.N.,
RA   Birkenmeyer L.G., Chalmers M.L., Pilot-Matias T.J., Dexai S.M.;
RT   "Molecular and biophysical characterization of TT virus: evidence for
RT   a new virus family infecting humans.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:3177-3182(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99350006; PubMed=10423143;
RA   Erker J.C., Leary T.P., Desai S.M., Chalmers M.L., Mushahwar I.K.;
RT   "Analyses of TT virus full-length genomic sequences.";
RL   J. Gen. Virol. 80:1743-1750(1999).
DR   EMBL; AF122914; AAD44680.1; -; Genomic_DNA.
DR   InterPro; IPR004118; TT_ORF2.
DR   Pfam; PF02957; TT_ORF2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   150 AA;  15621 MW;  43271527702CFE94 CRC64;

  Query Match             92.1%;  Score 35;  DB 2;  Length 150;
  Best Local Similarity   85.7%;  Pred. No. 2.8e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              ||||:||
Db         92 RALPAPP 98


Search completed: April  4, 2006, 13:15:15
Job time : 6.35079 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:07:57 ; Search time 0.888743 Seconds
(without alignments)
651.178 Million cell updates/sec

Title: US-10-632-388-293
Perfect score: 38
Sequence: 1 RALPSPP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	7	2	US-09-428-082B-293	Sequence 293, App
2	38	100.0	13	2	US-08-602-999A-78	Sequence 78, Appl
3	38	100.0	13	2	US-08-278-865-78	Sequence 78, Appl
4	38	100.0	13	2	US-09-500-124-78	Sequence 78, Appl
5	38	100.0	13	2	US-09-938-315-78	Sequence 78, Appl
6	38	100.0	25	2	US-08-278-865-45	Sequence 45, Appl
7	38	100.0	25	2	US-09-938-315-45	Sequence 45, Appl
8	38	100.0	26	2	US-08-602-999A-45	Sequence 45, Appl
9	38	100.0	26	2	US-09-500-124-45	Sequence 45, Appl
10	35	92.1	204	2	US-09-252-991A-16784	Sequence 16784, A
11	35	92.1	268	2	US-09-902-540-13693	Sequence 13693, A
12	35	92.1	485	2	US-10-214-269-20	Sequence 20, Appl
13	35	92.1	486	2	US-09-354-123-2	Sequence 2, Appli
14	35	92.1	610	2	US-09-949-016-7708	Sequence 7708, Ap
15	34	89.5	15	2	US-08-602-999A-333	Sequence 333, App
16	34	89.5	15	2	US-09-500-124-333	Sequence 333, App
17	34	89.5	115	2	US-09-252-991A-30248	Sequence 30248, A
18	34	89.5	476	2	US-10-214-269-2	Sequence 2, Appli
19	34	89.5	486	2	US-10-214-269-12	Sequence 12, Appl
20	34	89.5	500	2	US-09-354-123-6	Sequence 6, Appli
21	34	89.5	500	2	US-10-214-269-19	Sequence 19, Appl
22	34	89.5	502	2	US-10-214-269-16	Sequence 16, Appl
23	34	89.5	507	2	US-10-214-269-14	Sequence 14, Appl
24	33	86.8	7	1	US-08-230-047-35	Sequence 35, Appl
25	33	86.8	12	1	US-08-230-047-18	Sequence 18, Appl
26	33	86.8	67	2	US-09-248-796A-25591	Sequence 25591, A
27	33	86.8	92	2	US-09-270-767-61654	Sequence 61654, A
28	33	86.8	135	2	US-09-199-637A-217	Sequence 217, App
29	33	86.8	163	2	US-09-270-767-45418	Sequence 45418, A
30	33	86.8	167	2	US-09-949-016-10516	Sequence 10516, A
31	33	86.8	173	2	US-09-252-991A-19768	Sequence 19768, A
32	33	86.8	189	2	US-09-252-991A-19341	Sequence 19341, A
33	33	86.8	224	2	US-09-252-991A-24969	Sequence 24969, A
34	33	86.8	290	2	US-09-949-016-8166	Sequence 8166, Ap
35	33	86.8	295	2	US-09-199-637A-341	Sequence 341, App
36	33	86.8	351	2	US-09-245-041-11	Sequence 11, Appl
37	33	86.8	351	2	US-09-358-055B-11	Sequence 11, Appl
38	33	86.8	351	2	US-09-893-238-11	Sequence 11, Appl
39	33	86.8	369	1	US-08-230-047-5	Sequence 5, Appli
40	33	86.8	388	2	US-09-252-991A-30849	Sequence 30849, A
41	33	86.8	394	2	US-09-252-991A-27774	Sequence 27774, A
42	33	86.8	440	2	US-09-270-767-46101	Sequence 46101, A
43	33	86.8	469	2	US-09-830-230A-93	Sequence 93, Appl
44	33	86.8	495	2	US-09-252-991A-24229	Sequence 24229, A
45	33	86.8	516	2	US-09-252-991A-21329	Sequence 21329, A

ALIGNMENTS

RESULT 1
US-09-428-082B-293
; Sequence 293, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 293
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SH3 ANTAGONIST PEPTIDE
US-09-428-082B-293

Query Match 100.0%; Score 38; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALPSPP 7
|||||||
Db 1 RALPSPP 7

RESULT 2
US-08-602-999A-78
; Sequence 78, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:

```
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/602,999A
;      FILING DATE:  16-FEB-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-202
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  78:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-602-999A-78

  Query Match             100.0%;  Score 38;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db          4 RALPSPP 10


RESULT 3
US-08-278-865-78
; Sequence 78, Application US/08278865
; Patent No. 6303574
;  GENERAL INFORMATION:
;    APPLICANT:  KAY, BRIAN K.
;    APPLICANT:  SPARKS, ANDREW B.
;    APPLICANT:  THORN, JUDITH M.
;    APPLICANT:  QUILLIAM, LAWRENCE A.
;    APPLICANT:  DER, CHANNING J.
;    TITLE OF INVENTION:  Src SH3 BINDING PEPTIDES AND METHODS OF
;    TITLE OF INVENTION:  ISOLATING AND USING SAME
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/278,865
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Villacorta, Gilberto M.
;      REGISTRATION NUMBER:  34,038
;      REFERENCE/DOCKET NUMBER:  4980-007-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  78:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-278-865-78

  Query Match             100.0%;  Score 38;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db          4 RALPSPP 10


RESULT 4
US-09-500-124-78
; Sequence 78, Application US/09500124
; Patent No. 6432920
;  GENERAL INFORMATION:
;    APPLICANT:  SPARKS, Andrew B.
;    APPLICANT:  KAY, Brian K.
;    APPLICANT:  THORN, Judith M.
;    APPLICANT:  QUILLIAM, Lawrence A.
;    APPLICANT:  DER, Channing J.
;    APPLICANT:  FOWLKES, Dana M.
;    APPLICANT:  RIDER, James E.
;    TITLE OF INVENTION:  SH3 BINDING PEPTIDES AND METHODS OF
;    TITLE OF INVENTION:  ISOLATING AND USING SAME
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/500,124
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/602,999
;      FILING DATE:  16-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-202
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  78:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-09-500-124-78

  Query Match             100.0%;  Score 38;  DB 2;  Length 13;
```

```
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db          4 RALPSPP 10


RESULT 5
US-09-938-315-78
; Sequence 78, Application US/09938315
; Patent No. 6703482
;  GENERAL INFORMATION:
;       APPLICANT: KAY, BRIAN K.
;                  SPARKS, ANDREW B.
;                  THORN, JUDITH M.
;                  QUILLIAM, LAWRENCE A.
;                  DER, CHANNING J.
;       TITLE OF INVENTION: Src SH3 BINDING PEPTIDES AND METHODS OF
;                           ISOLATING AND USING SAME
;       NUMBER OF SEQUENCES: 106
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;                       P.C.
;            STREET: 1755 S. Jefferson Davis Highway, Suite 400
;            CITY: Arlington
;            STATE: Virginia
;            COUNTRY: U.S.A.
;            ZIP: 22202
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/938,315
;            FILING DATE: 23-Aug-2001
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Villacorta, Gilberto M.
;            REGISTRATION NUMBER: 34,038
;            REFERENCE/DOCKET NUMBER: 4980-007-0
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (703) 413-3000
;            TELEFAX: (703) 413-2220
;            TELEX: 248855 OPAT UR
;  INFORMATION FOR SEQ ID NO: 78:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 13 amino acids
;            TYPE: amino acid
;            TOPOLOGY: unknown
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 78:
US-09-938-315-78

  Query Match             100.0%;  Score 38;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db          4 RALPSPP 10


RESULT 6
US-08-278-865-45
; Sequence 45, Application US/08278865
; Patent No. 6303574
;  GENERAL INFORMATION:
;    APPLICANT:  KAY, BRIAN K.
;    APPLICANT:  SPARKS, ANDREW B.
;    APPLICANT:  THORN, JUDITH M.
;    APPLICANT:  QUILLIAM, LAWRENCE A.
;    APPLICANT:  DER, CHANNING J.
;    TITLE OF INVENTION:  Src SH3 BINDING PEPTIDES AND METHODS OF
;    TITLE OF INVENTION:  ISOLATING AND USING SAME
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/278,865
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Villacorta, Gilberto M.
;      REGISTRATION NUMBER:  34,038
;      REFERENCE/DOCKET NUMBER:  4980-007-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-278-865-45

  Query Match             100.0%;  Score 38;  DB 2;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 4.1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db         12 RALPSPP 18


RESULT 7
US-09-938-315-45
; Sequence 45, Application US/09938315
; Patent No. 6703482
;  GENERAL INFORMATION:
;       APPLICANT: KAY, BRIAN K.
;                  SPARKS, ANDREW B.
;                  THORN, JUDITH M.
;                  QUILLIAM, LAWRENCE A.
;                  DER, CHANNING J.
;       TITLE OF INVENTION: Src SH3 BINDING PEPTIDES AND METHODS OF
;                           ISOLATING AND USING SAME
;       NUMBER OF SEQUENCES: 106
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;                       P.C.
;            STREET: 1755 S. Jefferson Davis Highway, Suite 400
;            CITY: Arlington
;            STATE: Virginia
;            COUNTRY: U.S.A.
;            ZIP: 22202
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
```

```
;             APPLICATION NUMBER: US/09/938,315
;             FILING DATE: 23-Aug-2001
;             CLASSIFICATION: <Unknown>
;          ATTORNEY/AGENT INFORMATION:
;             NAME: Villacorta, Gilberto M.
;             REGISTRATION NUMBER: 34,038
;             REFERENCE/DOCKET NUMBER: 4980-007-0
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (703) 413-3000
;             TELEFAX: (703) 413-2220
;             TELEX: 248855 OPAT UR
;    INFORMATION FOR SEQ ID NO: 45:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 25 amino acids
;             TYPE: amino acid
;             TOPOLOGY: unknown
;          MOLECULE TYPE: peptide
;          SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-938-315-45

  Query Match             100.0%;  Score 38;  DB 2;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 4.1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db         12 RALPSPP 18


RESULT 8
US-08-602-999A-45
; Sequence 45, Application US/08602999A
; Patent No. 6184205
;  GENERAL INFORMATION:
;    APPLICANT:  SPARKS, Andrew B.
;    APPLICANT:  KAY, Brian K.
;    APPLICANT:  THORN, Judith M.
;    APPLICANT:  QUILLIAM, Lawrence A.
;    APPLICANT:  DER, Channing J.
;    APPLICANT:  FOWLKES, Dana M.
;    APPLICANT:  RIDER, James E.
;    TITLE OF INVENTION:  SH3 BINDING PEPTIDES AND METHODS OF
;    TITLE OF INVENTION:  ISOLATING AND USING SAME
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/602,999A
;      FILING DATE:  16-FEB-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-202
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-602-999A-45

  Query Match             100.0%;  Score 38;  DB 2;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 4.3;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db         13 RALPSPP 19


RESULT 9
US-09-500-124-45
; Sequence 45, Application US/09500124
; Patent No. 6432920
;  GENERAL INFORMATION:
;    APPLICANT:  SPARKS, Andrew B.
;    APPLICANT:  KAY, Brian K.
;    APPLICANT:  THORN, Judith M.
;    APPLICANT:  QUILLIAM, Lawrence A.
;    APPLICANT:  DER, Channing J.
;    APPLICANT:  FOWLKES, Dana M.
;    APPLICANT:  RIDER, James E.
;    TITLE OF INVENTION:  SH3 BINDING PEPTIDES AND METHODS OF
;    TITLE OF INVENTION:  ISOLATING AND USING SAME
;    NUMBER OF SEQUENCES:  467
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/500,124
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/602,999
;      FILING DATE:  16-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  1101-202
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-09-500-124-45

  Query Match             100.0%;  Score 38;  DB 2;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 4.3;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RALPSPP 7
              |||||||
Db         13 RALPSPP 19


RESULT 10
US-09-252-991A-16784
```

; Sequence 16784, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16784
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16784

Query Match 92.1%; Score 35; DB 2; Length 204;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RALPSPP 7
              ||:||||
Db         26 RAVPSPP 32
```

RESULT 11
US-09-902-540-13693
; Sequence 13693, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13693
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13693

Query Match 92.1%; Score 35; DB 2; Length 268;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RALPSPP 7
              ||||:||
Db        151 RALPAPP 157
```

RESULT 12
US-10-214-269-20
; Sequence 20, Application US/10214269
; Patent No. 6844485
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene H.
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Orozco, Jr., Emil M.
; APPLICANT: Rasco-Gaunt, Sonriza
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Phytochelatin Synthase
; FILE REFERENCE: BB1511 US NA
; CURRENT APPLICATION NUMBER: US/10/214,269
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/310,521
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-214-269-20

Query Match 92.1%; Score 35; DB 2; Length 485;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RALPSPP 7
              |:|||||
Db          9 RSLPSPP 15
```

RESULT 13
US-09-354-123-2
; Sequence 2, Application US/09354123
; Patent No. 6489537
; GENERAL INFORMATION:
; APPLICANT: Rea, Philip A.
; APPLICANT: Vatamaniuk, Olena K.
; APPLICANT: Mari, Stephane
; APPLICANT: Lu, Yu-Ping
; APPLICANT: Schroeder, Julian I.
; APPLICANT: Kim, Eugene J.
; APPLICANT: Clemens, Stephan
; TITLE OF INVENTION: NOVEL PHYTOCHELATIN SYNTHASES AND USES THEREFOR
; FILE REFERENCE: 9596-102U1/209596.0289
; CURRENT APPLICATION NUMBER: US/09/354,123
; CURRENT FILING DATE: 1999-07-15
; EARLIER APPLICATION NUMBER: 09/315,449
; EARLIER FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: 60/095,624
; EARLIER FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-354-123-2

Query Match 92.1%; Score 35; DB 2; Length 486;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RALPSPP 7
              |:|||||
Db          9 RSLPSPP 15
```

RESULT 14
US-09-949-016-7708
; Sequence 7708, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7708
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7708

Query Match 92.1%; Score 35; DB 2; Length 610;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RALPSPP 7
              ||||:||
Db         56 RALPAPP 62
```

RESULT 15
US-08-602-999A-333
; Sequence 333, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 333:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-333

Query Match 89.5%; Score 34; DB 2; Length 15;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RALPSPP 7
              |||| ||
Db          5 RALPGPP 11
```

Search completed: April 4, 2006, 13:09:43
Job time : 1.88874 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:06 ; Search time 3.47251 Seconds
(without alignments)
885.713 Million cell updates/sec

Title: US-10-632-388-294
Perfect score: 38
Sequence: 1 RRLPRTP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	7	3	AAB17238	Aab17238 SH3 antag
2	38	100.0	7	5	ABB73231	Abb73231 Src homol
3	38	100.0	7	7	ADJ73385	Adj73385 SH3 antag
4	38	100.0	7	8	ADJ53019	Adj53019 CH1 delet
5	38	100.0	7	8	ADJ51980	Adj51980 CH1 delet
6	38	100.0	13	2	AAW11122	Aaw11122 Src SH3 d
7	38	100.0	45	2	AAW16934	Aaw16934 Random re
8	38	100.0	45	2	AAW25497	Aaw25497 Random pe
9	38	100.0	269	4	AAU63824	Aau63824 Propionib
10	38	100.0	269	6	ABM60343	Abm60343 Propionib
11	38	100.0	289	6	ABU36530	Abu36530 Protein e
12	38	100.0	295	6	ABU34594	Abu34594 Protein e
13	38	100.0	5754	5	ABP62761	Abp62761 S. roseos
14	38	100.0	5830	7	ADJ72173	Adj72173 Streptomy
15	36	94.7	310	7	ABO72761	Abo72761 Pseudomon
16	35	92.1	875	7	ABO81310	Abo81310 Pseudomon
17	35	92.1	1023	4	ABB61882	Abb61882 Drosophil
18	35	92.1	1261	2	AAW75995	Aaw75995 GTPase ac
19	35	92.1	1261	3	AAY90268	Aay90268 Human GTP
20	35	92.1	1261	5	ADI17197	Adi17197 Human NOV
21	35	92.1	1261	8	ADL61268	Adl61268 Human tyr
22	34	89.5	1451	8	ADK71828	Adk71828 Human kin
23	34	89.5	1452	8	ADH09512	Adh09512 Human hos
24	34	89.5	1452	8	ADH09513	Adh09513 Human hos
25	34	89.5	1511	9	ADV97847	Adv97847 Murine pr
26	34	89.5	1512	8	ADH09514	Adh09514 Human hos
27	34	89.5	1512	8	ADH09515	Adh09515 Human hos
28	33	86.8	40	4	ABB41972	Abb41972 Peptide #
29	33	86.8	40	4	AAM35773	Aam35773 Peptide #
30	33	86.8	40	4	AAM75663	Aam75663 Human bon
31	33	86.8	40	4	AAM62850	Aam62850 Human bra
32	33	86.8	40	4	ABG57404	Abg57404 Human liv
33	33	86.8	40	5	ABG45166	Abg45166 Human pep
34	33	86.8	68	7	ADD71587	Add71587 Human uri
35	33	86.8	86	4	AAB46485	Aab46485 B. lichen
36	33	86.8	126	8	ADY23754	Ady23754 Plant ful
37	33	86.8	140	7	ABO78884	Abo78884 Pseudomon
38	33	86.8	141	7	ABO68120	Abo68120 Pseudomon
39	33	86.8	158	4	AAU43484	Aau43484 Propionib
40	33	86.8	158	6	ABM40003	Abm40003 Propionib
41	33	86.8	164	4	AAO07541	Aao07541 Human pol
42	33	86.8	181	8	ADX66532	Adx66532 Plant ful
43	33	86.8	183	4	ABG04951	Abg04951 Novel hum
44	33	86.8	198	8	ADX66082	Adx66082 Plant ful
45	33	86.8	210	8	ADX90372	Adx90372 Plant ful

ALIGNMENTS

RESULT 1
AAB17238
ID AAB17238 standard; peptide; 7 AA.
XX
AC AAB17238;
XX
DT 31-OCT-2000 (first entry)
XX
DE SH3 antagonist peptide sequence SEQ ID NO:294.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 39; Page 299; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

```
CC   c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC   of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC   thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC   cells from the present invention can be used for producing pharmaceutical
CC   compositions. The compositions are useful for treating cancer, asthma,
CC   thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC   a Fab domain) can provide a longer half-life or incorporate functions
CC   such as Fc receptor binding, protein A binding, complement fixation, and
CC   possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC   AAB18003 represent nucleotide and amino acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 38;  DB 3;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db          1 RRLPRTP 7


RESULT 2
ABB73231
ID   ABB73231 standard; peptide; 7 AA.
XX
AC   ABB73231;
XX
DT   05-APR-2002  (first entry)
XX
DE   Src homology3 (SH3) antagonist peptide SEQ ID NO:294.
XX
KW   Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW   erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW   TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW   TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW   MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW   cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW   antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW   neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW   cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW   sleep disorder; neurological degenerative disease; anaemia;
KW   thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW   Fanconi's syndrome.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200183525-A2.
XX
PD   08-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-US014310.
XX
PR   03-MAY-2000; 2000US-00563286.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham JC,  Boone TC,  Gudas JM;
XX
DR   WPI; 2002-130313/17.
XX
PT   Novel vehicle-peptide molecule or its multimers useful for treating
PT   inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT   diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS   Claim 39; Page 55; 176pp; English.
XX
CC   The present invention describes a vehicle-peptide molecule (I) or its
CC   multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC   cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC   antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC   neuroprotective activities. (I) can be used as a therapeutic or
CC   prophylactic agent as well as for screening purposes. (I) is useful for
CC   diagnosing diseases characterised by dysfunction of their associated
CC   protein of interest, for identifying normal or abnormal proteins of
CC   interest, as a part of diagnostic kit to detect the presence of their
CC   proteins of interest in a biological sample. Additionally, (I) is useful
CC   for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC   rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC   infertility, and neurological degenerative diseases. (I), comprising EPO-
CC   mimetic compounds are useful for treating disorders characterised by low
CC   red blood cell levels such as anaemia. The TPO-mimetic comprising
CC   compounds are useful for treating conditions that involve an existing
CC   megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC   deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC   tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC   and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC   represent amino acid and nucleic acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 38;  DB 5;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db          1 RRLPRTP 7


RESULT 3
ADJ73385
ID   ADJ73385 standard; peptide; 7 AA.
XX
AC   ADJ73385;
XX
DT   06-MAY-2004  (first entry)
XX
DE   SH3 antagonist peptide sequence SeqID 840.
XX
KW   mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW   cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW   immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW   SH3.
XX
OS   Synthetic.
XX
PN   WO2003084477-A2.
XX
PD   16-OCT-2003.
XX
PF   24-MAR-2003; 2003WO-US009139.
XX
PR   29-MAR-2002; 2002US-0368791P.
XX
PA   (CENZ ) CENTOCOR INC.
XX
PI   Heavner GA,  Knight DM,  Scallon BJ,  Ghrayeb J;
XX
DR   WPI; 2003-804237/75.
XX
PT   New CDR mimetibody comprising a portion of a heavy or light chain
PT   variable region comprising human framework or ligand binding region,
PT   useful for preparing a composition for treating e.g., immune,
PT   cardiovascular or neurologic disease.
XX
PS   Disclosure; SEQ ID NO 840; 97pp; English.
XX
CC   This invention relates to novel mammalian CDR mimetibodies, specific
CC   portions or variants thereof. Specifically, it refers to an antibody
CC   fragment where a protein has been inserted into, or replaces a portion
CC   of, one or more CDR regions, such that each CDR mimetibody comprises at
CC   least one portion of a heavy chain or light chain variable region, which
```

CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is an SH3 antagonist peptide sequence used to make a
CC mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLPRTP 7
|||||||
Db 1 RRLPRTP 7

RESULT 4
ADJ53019
ID ADJ53019 standard; peptide; 7 AA.
XX
AC ADJ53019;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID840.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Claim 3; SEQ ID NO 840; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 38; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLPRTP 7
|||||||
Db 1 RRLPRTP 7

RESULT 5
ADJ51980
ID ADJ51980 standard; peptide; 7 AA.
XX
AC ADJ51980;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID840.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 15; SEQ ID NO 840; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be

```
CC   useful for the development of compounds with an osteopathic,
CC   cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC   gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC   immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC   antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC   respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC   modulator or cytokine-agonist. The methods and compositions of the
CC   present invention are useful for the diagnosis, prevention and/or
CC   treatment of diseases or conditions associated with aberrant expression
CC   or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC   cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC   endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC   obstetric, haematologic, immunological, allergic, infectious,
CC   musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC   pediatric, psychiatric, renal or pulmonary disorders. The present
CC   sequence is that of a peptide which may be used during the creation of a
CC   mimetibody of the invention.
XX
SQ   Sequence 7 AA;

  Query Match             100.0%;  Score 38;  DB 8;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db          1 RRLPRTP 7


RESULT 6
AAW11122
ID   AAW11122 standard; peptide; 13 AA.
XX
AC   AAW11122;
XX
DT   25-JUN-1997  (first entry)
XX
DE   Src SH3 domain-binding peptide used in signal transduction modulation.
XX
KW   Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW   protein tyrosine kinase; signal transduction; RNA processing;
KW   trafficking; translation.
XX
OS   Synthetic.
XX
PN   WO9603649-A1.
XX
PD   08-FEB-1996.
XX
PF   24-JUL-1995;   95WO-US009382.
XX
PR   22-JUL-1994;   94US-00278865.
PR   07-JUN-1995;   95US-00483555.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ;
XX
DR   WPI; 1996-117151/12.
XX
PT   Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT   proteins - useful for e.g. modulating signal transduction pathways at the
PT   cellular level, esp. protein tyrosine kinase-mediated.
XX
PS   Claim 40; Page 84; 116pp; English.
XX
CC   AAW11098-W11124 are peptides that bind to the Src SH3 domain. The SH3
CC   binding peptides are useful in modulating signal transduction pathways at
CC   the cellular level (especially protein tyrosine kinase-mediated),
CC   modulating oncogenic protein activity, or providing compounds for the
CC   development of drugs with the ability to modulate broad classes, as well
CC   as specific classes, of proteins involved in signal transduction and also
CC   for regulating the processing, trafficking or translation of RNA.
CC   Conjugates of the peptides with detectable labels or imaging agents are
CC   useful for imaging cells, tissues and organs in which Src or Src-related
CC   proteins are expressed
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 38;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db          4 RRLPRTP 10


RESULT 7
AAW16934
ID   AAW16934 standard; peptide; 45 AA.
XX
AC   AAW16934;
XX
DT   27-JUN-1997  (first entry)
XX
DE   Random recombinant SH3 domain binding peptide.
XX
KW   Src; SH3; Src homology region 3; binding affinity; oncogenic protein;
KW   protein tyrosine kinase; signal transduction; RNA processing;
KW   trafficking; translation.
XX
OS   Synthetic.
XX
PN   WO9603649-A1.
XX
PD   08-FEB-1996.
XX
PF   24-JUL-1995;   95WO-US009382.
XX
PR   22-JUL-1994;   94US-00278865.
PR   07-JUN-1995;   95US-00483555.
XX
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Sparks AB,  Kay BK,  Thorn JM,  Quilliam LA,  Der CJ;
XX
DR   WPI; 1996-117151/12.
XX
PT   Peptide with binding affinity for Src homology region 3 (SH3) domains of
PT   proteins - useful for e.g. modulating signal transduction pathways at the
PT   cellular level, esp. protein tyrosine kinase-mediated.
XX
PS   Disclosure; Fig 1; 116pp; English.
XX
CC   AAW16924-W16948 are random recombinant peptides derived from one of three
CC   peptide libraries, T9, T12 and R8C. The peptides are all SH3 domain-
CC   binding peptides. SH3 binding peptides are useful in modulating signal
CC   transduction pathways at the cellular level (especially protein tyrosine
CC   kinase-mediated), modulating oncogenic protein activity, or providing
CC   compounds for the development of drugs with the ability to modulate broad
CC   classes, as well as specific classes, of proteins involved in signal
CC   transduction and also for regulating the processing, trafficking or
CC   translation of RNA. Conjugates of the peptides with detectable labels or
CC   imaging agents are useful for imaging cells, tissues and organs in which
CC   Src or Src-related proteins are expressed
XX
SQ   Sequence 45 AA;

  Query Match             100.0%;  Score 38;  DB 2;  Length 45;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db         31 RRLPRTP 37
```

RESULT 8
AAW25497
ID AAW25497 standard; peptide; 45 AA.
XX
AC AAW25497;
XX
DT 27-MAR-1998 (first entry)
XX
DE Random peptide recombinant clone T9.SRC3.1.
XX
KW Cortactin; SH3 domain; binding peptide; Src homology region 3;
KW tyrosine kinase; immune response; lymphokine; interleukin 1; Nck; Abl;
KW PLCgamma; p53bp2; Crk; Yes; Grb2.
XX
OS Synthetic.
OS Unidentified.
XX
PN WO9730074-A1.
XX
PD 21-AUG-1997.
XX
PF 14-FEB-1997; 97WO-US002298.
XX
PR 16-FEB-1996; 96US-00602999.
XX
PA (CYTO-) CYTOGEN CORP.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Sparks AB, Kay BK, Thorn JM, Quilliam LA, Der CJ, Fowlkes DM;
PI Rider JE;
XX
DR WPI; 1997-424972/39.
XX
PT Src homology region 3 binding peptide - used to activate Src tyrosine
PT kinase(s) and to stimulate immune response by increasing production of
PT certain lymphokine(s), e.g. interleukin-1.
XX
PS Disclosure; Fig 5; 131pp; English.
XX
CC The present sequence represents a random peptide recombinant isolated by
CC the method of the present invention. SH3 (Src homology region 3) binding
CC peptides are selected from: (a) peptides which bind the SH3 domain of
CC Cortactin; (b) peptides which bind the middle SH3 domain of Nck; (c)
CC peptides which bind the SH3 domain of Abl; (d) peptides which bind the
CC SH3 domain of Src; (e) peptides which bind the SH3 domain of PLC gamma;
CC (f) peptides which bind the SH3 domain of p53bp2; (g) peptides which bind
CC the amino-terminal SH3 domain of Crk; (h) peptides which bind the SH3
CC domain of Yes; and (i) peptides which bind the amino-terminal SH3 domain
CC of Grb2. The purified binding peptides can be used in the method to
CC identify inhibitors of their binding to their respective SH3 domains,
CC which could be used to modulate the pharmacological activity of proteins
CC or polypeptide containing the SH3 domain. The peptides can also be used
CC to activate Src or Src-related protein tyrosine kinases, to stimulate the
CC immune response by increasing the production of certain lymphokines, e.g.
CC tumour necrosis factor-alpha and interleukin-1, or to deliver a
CC conjugated molecule to certain cellular compartments containing Src or
CC Src related proteins
XX
SQ Sequence 45 AA;

```
  Query Match           100.0%;  Score 38;  DB 2;  Length 45;
  Best Local Similarity  100.0%;  Pred. No. 16;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db         31 RRLPRTP 37
```

RESULT 9
AAU63824
ID AAU63824 standard; protein; 269 AA.
XX
AC AAU63824;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #24720.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59636.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 25019; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 269 AA;

```
  Query Match           100.0%;  Score 38;  DB 4;  Length 269;
  Best Local Similarity  100.0%;  Pred. No. 80;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        127 RRLPRTP 133
```

RESULT 10
ABM60343

```
ID   ABM60343 standard; protein; 269 AA.
XX
AC   ABM60343;
XX
DT   20-OCT-2003  (first entry)
XX
DE   Propionibacterium acnes predicted ORF-encoded polypeptide #25019.
XX
KW   Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW   immunostimulant; immune response; vaccine.
XX
OS   Propionibacterium acnes.
XX
PN   WO2003033515-A1.
XX
PD   24-APR-2003.
XX
PF   11-OCT-2002; 2002WO-US032727.
XX
PR   15-OCT-2001; 2001US-00978825.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Mitcham JL,  Skeiky YAW,  Persing DH,  Bhatia A,  Maisonneuve JL;
PI   Zhang Y,  Wang S,  Jen S,  Lodes MJ,  Benson DR,  Jones R,  Carter D;
PI   Barth B,  Vallieve-Douglass J;
XX
DR   WPI; 2003-381789/36.
DR   N-PSDB; ACF64565.
XX
PT   New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT   polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT   or for stimulating an immune response specific for a P. acnes protein.
XX
PS   Example 1; SEQ ID NO 25019; 1481pp; English.
XX
CC   The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC   encoding a Propionibacterium acnes protein. The invention also relates to
CC   polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC   immunogenic fragments of P. acnes polypeptides. The invention
CC   additionally encompasses expression vectors and host cells comprising a
CC   polynucleotide of the invention; antibodies against polypeptides of the
CC   invention; fusion proteins comprising a polypeptide of the invention; a
CC   method for stimulating an immune response specific for a P. acnes
CC   polypeptide and an isolated T cell population comprising T cells prepared
CC   via this method; a vaccine composition (comprising P. acnes polypeptides,
CC   polynucleotides, antibodies, fusion proteins, T cell populations, or
CC   antigen-presenting cells that express the polypeptide); a method and kit
CC   for detecting or determining the presence or absence of P. acnes in a
CC   patient; and a method for inhibiting the development of P. acnes in a
CC   patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC   proteins, T cell populations or antigen-presenting cells that express the
CC   polypeptides are useful for diagnosing, preventing or treating acne
CC   vulgaris, or for stimulating an immune response specific for a P. acnes
CC   protein. The polynucleotides can also be used as probes or primers for
CC   nucleic acid hybridisation. The vaccine composition is useful for the
CC   stimulation of an immune response against P. acnes, or for treating acne,
CC   and the kit is useful for performing a diagnostic assay. The present
CC   sequence represents a polypeptide predicted to be encoded by an ORF (open
CC   reading frame) contained within the P. acnes polynucleotides of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 269 AA;

  Query Match             100.0%;  Score 38;  DB 6;  Length 269;
  Best Local Similarity   100.0%;  Pred. No. 80;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        127 RRLPRTP 133
```

```
RESULT 11
ABU36530
ID   ABU36530 standard; protein; 289 AA.
XX
AC   ABU36530;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #22057.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA40400.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 64454; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
```

```
XX
SQ   Sequence 289 AA;

  Query Match             100.0%;  Score 38;  DB 6;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 85;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        266 RRLPRTP 272


RESULT 12
ABU34594
ID   ABU34594 standard; protein; 295 AA.
XX
AC   ABU34594;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #20121.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Mycobacterium bovis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA38464.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 62518; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 295 AA;

  Query Match             100.0%;  Score 38;  DB 6;  Length 295;
  Best Local Similarity   100.0%;  Pred. No. 87;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        272 RRLPRTP 278


RESULT 13
ABP62761
ID   ABP62761 standard; protein; 5754 AA.
XX
AC   ABP62761;
XX
DT   23-OCT-2002  (first entry)
XX
DE   S. roseosporus daptomycin non-ribosomal peptide synthetase DptA.
XX
KW   Daptomycin biosynthetic gene cluster; thioesterase; antibacterial;
KW   fungicide; virucide; antiparasitic; immunomodulator; antilipemic;
KW   cytostatic; gene therapy; antimitotic; immunomodulatory; siderophore;
KW   anti-cholesterolemic; agrochemical; non-ribosomal peptide synthetase;
KW   NRPS; DptA.
XX
OS   Streptomyces roseosporus.
XX
PN   WO200259322-A2.
XX
PD   01-AUG-2002.
XX
PF   17-OCT-2001; 2001WO-US032354.
XX
PR   17-OCT-2000; 2000US-0240879P.
PR   28-FEB-2001; 2001US-0272207P.
PR   06-AUG-2001; 2001US-0310385P.
XX
PA   (MIAO/) MIAO V P W.
PA   (BRIA/) BRIAN P.
PA   (BALT/) BALTZ R H.
PA   (SILV/) SILVA C J.
XX
PI   Miao VPW,  Brian P,  Baltz RH,  Silva CJ;
XX
DR   WPI; 2002-599794/64.
XX
PT   Isolated nucleic acid molecule from a bacterial daptomycin biosynthetic
PT   gene cluster encoding a thioesterase or thioesterase domain, useful for
PT   generating novel linear and cyclic peptides, and products in a cell.
XX
PS   Claim 25; Page 166-169; 227pp; English.
XX
CC   The invention relates to a novel isolated nucleic acid molecule
CC   comprising a sequence that encodes a thioesterase or thioesterase domain,
CC   derived from a bacterial daptomycin biosynthetic gene cluster. The
CC   proteins of the invention have antibacterial, fungicide, virucide,
CC   antiparasitic, immunomodulator, antilipemic, and cytostatic activity. The
CC   polynucleotides may have a use in gene therapy. The compositions and
CC   methods of the present invention are useful for generating novel linear
CC   and cyclic peptides and improving yield of a product in a cell expressing
```

```
CC   an daptomycin non-ribosomal peptide synthetase (NRPS) to be used as new
CC   compounds or in producing new compounds, such as antibiotics,
CC   antifungals, antivirals, antiparasitics, antimitotics, antitumour agents,
CC   immunomodulatory agents, anti-cholesterolemic agents, siderophores,
CC   agrochemicals and cytostatics. The sequence represents a S. roseosporus
CC   daptomycin non-ribosomal peptide synthetase of the invention
XX
SQ   Sequence 5754 AA;

  Query Match             100.0%;  Score 38;  DB 5;  Length 5754;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db       4579 RRLPRTP 4585


RESULT 14
ADJ72173
ID   ADJ72173 standard; protein; 5830 AA.
XX
AC   ADJ72173;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Streptomyces roseosporus DptA protein.
XX
KW   antibacterial; gene therapy; daptomycin biosynthesis gene cluster;
KW   daptomycin non-ribosomal peptide synthetase; DptBC;
KW   gram-positive bacterial infection.
XX
OS   Streptomyces roseosporus.
XX
PN   WO2003014297-A2.
XX
PD   20-FEB-2003.
XX
PF   31-JUL-2002; 2002WO-US024310.
XX
PR   06-AUG-2001; 2001US-0310385P.
PR   17-OCT-2001; 2001WO-US032354.
PR   10-MAY-2002; 2002US-0379866P.
XX
PA   (CUBI-) CUBIST PHARM INC.
XX
PI   Miao VPW,  Brian P,  Baltz RH,  Coeffet-Legal MF;
XX
DR   WPI; 2003-268192/26.
DR   N-PSDB; ADJ72363.
XX
PT   New isolated nucleic acid molecule encoding a daptomycin non-ribosomal
PT   peptide synthetase, useful for treatment of a gram-positive bacterial
PT   infection of skeletal muscle, skin, bloodstream, kidneys, heart, lung and
PT   bone.
XX
PS   Disclosure; SEQ ID NO 9; 292pp; English.
XX
CC   The invention relates to new isolated nucleic acid (NA) molecules from
CC   the Streptomyces roseosporus daptomycin biosynthesis gene cluster,
CC   especially a daptomycin non-ribosomal peptide synthetase (NRPS) or its
CC   subunit, where the (NA) molecule encodes DptBC, and is not pRHB159. The
CC   methods and compositions of the present invention are useful for
CC   treatment of a gram-positive bacterial infection of any organ or tissue
CC   in the body, including skeletal muscle, skin, bloodstream, kidneys,
CC   heart, lung and bone. This sequence represents the daptomycin
CC   biosynthesis protein DptA.
XX
SQ   Sequence 5830 AA;

  Query Match             100.0%;  Score 38;  DB 7;  Length 5830;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db       4655 RRLPRTP 4661


RESULT 15
ABO72761
ID   ABO72761 standard; protein; 310 AA.
XX
AC   ABO72761;
XX
DT   29-JUL-2004  (first entry)
XX
DE   Pseudomonas aeruginosa polypeptide #4936.
XX
KW   Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS   Pseudomonas aeruginosa.
XX
PN   US6551795-B1.
XX
PD   22-APR-2003.
XX
PF   18-FEB-1999;   99US-00252991.
XX
PR   18-FEB-1998;   98US-0074788P.
PR   27-JUL-1998;   98US-0094190P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Rubenfield MJ,  Nolling J,  Deloughery C,  Bush D;
XX
DR   WPI; 2003-615309/58.
DR   N-PSDB; ABD06332.
XX
PT   Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT   useful as molecular targets for diagnostics, prophylaxis and treatment of
PT   pathological conditions resulting from bacterial infection.
XX
PS   Disclosure; SEQ ID NO 21507; 455pp; English.
XX
CC   The invention relates to Pseudomonas aeruginosa polypeptides and the
CC   polynucleotides encoding them. The sequences are useful in diagnosis and
CC   therapy of pathological conditions, as molecular targets for diagnostics,
CC   prophylaxis and treatment of pathological conditions resulting from a
CC   bacterial infection, for evaluating a compound, such as a polypeptide,
CC   for the ability to bind a P. aeruginosa nucleic acid, as components of
CC   effective antibacterial targets, as targets for antibacterial drugs,
CC   including anti-P. aeruginosa drugs, as templates for recombinant
CC   production of P. aeruginosa-derived peptides or polypeptides, as target
CC   components for diagnosis and/or treatment of P. aeruginosa-caused
CC   infection, and in detection of P. aeruginosa sequences or other sequences
CC   of Pseudomonas species using biochip technology. Sequences ABO67826-
CC   ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format from USPTO at
CC   seqdata.uspto.gov/sequence.html
XX
SQ   Sequence 310 AA;

  Query Match             94.7%;  Score 36;  DB 7;  Length 310;
  Best Local Similarity   85.7%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              ||:||||
Db        230 RRMPRTP 236


Search completed: April  4, 2006, 13:07:35
Job time : 5.47251 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:05:37 ; Search time 1.14529 Seconds
(without alignments)
588.077 Million cell updates/sec

Title: US-10-632-388-294
Perfect score: 38
Sequence: 1 RRLPRTP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	289	2	D70809	probable pabC prot
2	35	92.1	761	2	T24230	hypothetical prote
3	35	92.1	1261	2	E59430	PTPL1-associated R
4	35	92.1	1335	2	JQ1258	RNA-directed RNA p
5	34	89.5	411	2	A84365	hypothetical prote
6	34	89.5	2326	2	B47447	calcium channel pr
7	33	86.8	98	2	S01566	hypothetical prote
8	33	86.8	281	2	F82832	pantoate-beta-alan
9	33	86.8	298	2	H83163	probable transcrip
10	33	86.8	324	2	A44241	clavaminate syntha
11	33	86.8	496	2	T09936	hypothetical prote
12	33	86.8	499	2	T34328	hypothetical prote
13	33	86.8	543	2	S62456	probable serine-th
14	33	86.8	573	2	T25397	hypothetical prote
15	33	86.8	650	2	G83465	conserved hypothet
16	33	86.8	657	2	JC7767	isoamylase (EC 3.2
17	33	86.8	677	2	H84382	heterodisulfide re
18	33	86.8	689	2	T35882	hypothetical prote
19	33	86.8	705	2	A70669	probable acid-CoA
20	33	86.8	810	2	T48835	lethal(2)denticlel
21	33	86.8	1088	2	S50925	hypothetical prote
22	33	86.8	5255	2	T31677	bacitracin synthet
23	32	84.2	109	2	A72546	hypothetical prote
24	32	84.2	111	2	S23601	hypothetical prote
25	32	84.2	267	2	B83109	probable transcrip
26	32	84.2	332	2	B88042	protein F56D12.3 [
27	32	84.2	511	2	D70507	hypothetical prote
28	32	84.2	583	2	C84788	probable myosin he
29	32	84.2	772	2	T32911	hypothetical prote
30	32	84.2	1132	2	T03844	telomerase catalyt
31	32	84.2	1334	2	T41524	rho1 gdp-gtp excha
32	31	81.6	303	2	D70955	hypothetical prote
33	31	81.6	304	2	T29421	hypothetical prote
34	31	81.6	385	2	E83506	probable MFS trans
35	31	81.6	387	2	H72607	hypothetical prote
36	31	81.6	408	2	B84518	hypothetical prote
37	31	81.6	411	2	T04987	hypothetical prote
38	31	81.6	487	2	T34887	probable transposa
39	31	81.6	504	2	A23282	RAD52 protein - ye
40	31	81.6	608	2	G75561	ABC transporter, A
41	31	81.6	750	1	JDVLVH	DNA-directed DNA p
42	31	81.6	761	2	JC5759	brain-specific ser
43	31	81.6	822	2	AG1911	hypothetical prote
44	31	81.6	825	2	T13473	DNA-directed DNA p
45	31	81.6	827	2	T13468	DNA-directed DNA p

ALIGNMENTS

RESULT 1
D70809
probable pabC protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70809
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genom
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70809
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-289 <COL>
A;Cross-references: UNIPARC:UPI00000D3AF9; GB:AL022004; GB:AL123456; NID:g3261550; PID
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: pabC

Query Match 100.0%; Score 38; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              |||||||
Db        266 RRLPRTP 272
```

RESULT 2
T24230
hypothetical protein R166.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24230
R;Matthews, P.
submitted to the EMBL Data Library, August 1995
A;Reference number: Z19859
A;Accession: T24230
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-761 <WIL>
A;Cross-references: UNIPROT:Q22005; UNIPARC:UPI000007C639; EMBL:Z50795; PIDN:CAA90665.
A;Experimental source: clone R166
C;Genetics:
A;Gene: CESP:R166.5
A;Map position: 2
A;Introns: 41/1; 198/1; 427/2; 535/2; 613/2; 747/3

Query Match 92.1%; Score 35; DB 2; Length 761;

Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              ||:||||
Db        180 RRVPRTP 186
```

RESULT 3
E59430
PTPL1-associated RhoGAP protein 1 [imported] - human
C;Species: Homo sapiens (man)
C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C;Accession: E59430
R;Saras, J.; Franzen, P.; Aspenstrom, P.; Hellman, U.; Gonez, L.J.; Heldin, C.-H.
submitted to GenBank, December 1997
A;Description: Homo sapiens PTPL1-associated RhoGAP 1 (PARG1), mRNA.
A;Reference number: E59430
A;Accession: E59430
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1261 <SAR>
A;Cross-references: UNIPROT:O15463; UNIPARC:UPI000007289E; GB:NP_004806; PID:g4758882; P

Query Match 92.1%; Score 35; DB 2; Length 1261;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              |:|||||
Db        565 RKLPRTP 571
```

RESULT 4
JQ1258
RNA-directed RNA polymerase (EC 2.7.7.48) - foxtail mosaic virus
N;Alternate names: 152.3K protein
C;Species: foxtail mosaic virus
C;Date: 05-Mar-1993 #sequence_revision 05-Mar-1993 #text_change 09-Jul-2004
C;Accession: JQ1258
R;Bancroft, J.B.; Rouleau, M.; Johnston, R.; Prins, L.; Mackie, G.A.
J. Gen. Virol. 72, 2173-2181, 1991
A;Title: The entire nucleotide sequence of foxtail mosaic virus RNA.
A;Reference number: JQ1258; MUID:91374015; PMID:1840610
A;Accession: JQ1258
A;Molecule type: genomic RNA
A;Residues: 1-1335 <BAN>
A;Cross-references: UNIPROT:P22168; UNIPARC:UPI0000134B1E; GB:M62730; NID:g325391; PIDN:
C;Superfamily: eggplant mosaic virus RNA-directed RNA polymerase
C;Keywords: nucleotidyltransferase; RNA biosynthesis

Query Match 92.1%; Score 35; DB 2; Length 1335;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              ||||:||
Db        526 RRLPKTP 532
```

RESULT 5
A84365
hypothetical protein Vng2148h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: A84365
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: A84365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <STO>
A;Cross-references: UNIPROT:Q9HND5; UNIPARC:UPI0000063A68; GB:AE004437; NID:g10581564;
C;Genetics:
A;Gene: VNG2148H

Query Match 89.5%; Score 34; DB 2; Length 411;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              ||||:||
Db        305 RRLPQTP 311
```

RESULT 6
B47447
calcium channel protein alpha-1 chain (variant doe-4) - electric ray (Discopyge ommata)
C;Species: Discopyge ommata
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B47447
R;Horne, W.A.; Ellinor, P.T.; Inman, I.; Zhou, M.; Tsien, R.W.; Schwarz, T.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 3787-3791, 1993
A;Title: Molecular diversity of Ca(2+) channel alpha 1 subunits from the marine ray Dis
A;Reference number: A47447; MUID:93248175; PMID:7683405
A;Accession: B47447
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-2326 <HOR>
A;Cross-references: UNIPROT:P56698; UNIPARC:UPI0000127266
A;Note: sequence extracted from NCBI backbone (NCBIP:130673)
C;Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 89.5%; Score 34; DB 2; Length 2326;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              |:|||||
Db       2165 RQLPRTP 2171
```

RESULT 7
S01566
hypothetical protein - human cytomegalovirus
C;Species: human cytomegalovirus, human herpesvirus 5
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S01566; S09782
R;Beck, S.; Barrell, B.G.
Nature 331, 269-272, 1988
A;Title: Human cytomegalovirus encodes a glycoprotein homologous to MHC class-I antigen
A;Reference number: S00661; MUID:88094735; PMID:2827039
A;Accession: S01566
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-98 <BEC>
A;Cross-references: UNIPROT:P16723; UNIPARC:UPI0000137B64; EMBL:Y00293
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A;Title: Analysis of the protein-coding content of the sequence of human cytomegaloviru
A;Reference number: S09749; MUID:90269039; PMID:2161319
A;Accession: S09782
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-98 <CHE>
A;Cross-references: UNIPARC:UPI0000137B64; EMBL:X17403; NID:g59591; PIDN:CAA35418.1; PI
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1989

Query Match 86.8%; Score 33; DB 2; Length 98;
Best Local Similarity 85.7%; Pred. No. 15;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              ||||| |
Db         12 RRLPRAP 18
```

RESULT 8
F82832
pantoate-beta-alanine ligase XF0230 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: F82832
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: F82832
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-281 <SIM>
A;Cross-references: UNIPROT:Q9PGR8; UNIPARC:UPI00001312AF; GB:AE003876; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0230
C;Superfamily: pantoate-beta-alanine ligase

Query Match 86.8%; Score 33; DB 2; Length 281;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              || |||| 
Db         70 RRYPRTP 76
```

RESULT 9
H83163
probable transcription regulator PA3845 [imported] - Pseudomonas aeruginosa (strain PAO1
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: H83163
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83163
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-298 <STO>
A;Cross-references: UNIPROT:Q9HXG1; UNIPARC:UPI00000C5B3A; GB:AE004802; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA3845
C;Superfamily: regulatory protein ampR

Query Match 86.8%; Score 33; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RLPRTP 7
              ||||||
Db        176 RLPRTP 181
```

RESULT 10
A44241
clavaminate synthase 1 - Streptomyces clavuligerus
C;Species: Streptomyces clavuligerus
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Oct-2004
C;Accession: A44241
R;Marsh, E.N.; Chang, M.D.; Townsend, C.A.
Biochemistry 31, 12648-12657, 1992
A;Title: Two isozymes of clavaminate synthase central to clavulanic acid formation: cl
A;Reference number: A44241; MUID:93112607; PMID:1472501
A;Accession: A44241
A;Status: preliminary
A;Molecule type: DNA; protein
A;Residues: 1-324 <MAR>
A;Cross-references: UNIPROT:Q05581; UNIPARC:UPI000003064B; GB:L06213; NID:g153219; PID
A;Note: sequence extracted from NCBI backbone (NCBIN:121675, NCBIP:121676)
C;Superfamily: clavaminate synthase 1

Query Match 86.8%; Score 33; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RLPRTP 7
              ||||||
Db         19 RLPRTP 24
```

RESULT 11
T09936
hypothetical protein T16L4.240 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T09936
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16897
A;Accession: T09936
A;Molecule type: DNA
A;Residues: 1-496 <BEV>
A;Cross-references: UNIPROT:Q9SU78; UNIPARC:UPI00000A2401; EMBL:AL079344; GSPDB:GN0006
A;Experimental source: cultivar Columbia; BAC clone T16L4
C;Genetics:
A;Gene: ATSP:T16L4.240
A;Map position: 4
A;Introns: 85/2; 105/3; 134/3; 155/3; 177/3; 202/3; 228/1; 293/2; 320/3; 345/2; 382/3;

Query Match 86.8%; Score 33; DB 2; Length 496;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RRLPRTP 7
              || ||||
Db         15 RRAPRTP 21
```

RESULT 12
T34328
hypothetical protein K03E6.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T34328
R;Latreille, P.; Gattung, S.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of C. elegans cosmid K03E6.

A;Reference number: Z21506
A;Accession: T34328
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-499 <LAT>
A;Cross-references: UNIPROT:Q21194; UNIPARC:UPI000017BA86; EMBL:U55375; PIDN:AAC69046.1;
A;Experimental source: strain Bristol N2; clone K03E6
C;Genetics:
A;Gene: CESP:K03E6.7
A;Map position: X
A;Introns: 17/1; 68/2; 101/3; 149/3; 194/2; 302/3; 355/2; 406/3; 446/1

Query Match 86.8%; Score 33; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLPRTP 7
||||||
Db 440 RLPRTP 445

RESULT 13
S62456
probable serine-threonine-protein kinase - fission yeast (Schizosaccharomyces pombe) (fr
C;Species: Schizosaccharomyces pombe
C;Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 05-Oct-2004
C;Accession: S62456; T38567
R;Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, October 1995
A;Reference number: S62445
A;Accession: S62456
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-543 <BAD>
A;Cross-references: UNIPARC:UPI0000169030; EMBL:Z54354; NID:g1019398; PIDN:CAA91166.1; P
R;Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z21745
A;Accession: T38567
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-543 <BA2>
A;Cross-references: UNIPARC:UPI0000169030; EMBL:Z54354; PIDN:CAA91166.1; GSPDB:GN00066;
A;Experimental source: strain 972h-; cosmid c2G11
C;Genetics:
A;Gene: SPDB:SPAC2G11.01
A;Map position: 1L
C;Keywords: ATP
F;223-521/Domain: protein kinase homology <KIN>
F;231-239/Region: protein kinase ATP-binding motif

Query Match 86.8%; Score 33; DB 2; Length 543;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RRLPRTP 7
||||| |
Db 72 RRLPRAP 78

RESULT 14
T25397
hypothetical protein T28B11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25397
R;Kelly, P.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z20028
A;Accession: T25397
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-573 <WIL>
A;Cross-references: UNIPROT:Q22842; UNIPARC:UPI0000079B54; EMBL:Z73977; PIDN:CAA98290.1
A;Experimental source: clone T28B11
C;Genetics:
A;Gene: CESP:T28B11.1
A;Map position: 5
A;Introns: 32/2; 477/2; 539/1
C;Superfamily: Caenorhabditis elegans hypothetical protein T28B11.1

Query Match 86.8%; Score 33; DB 2; Length 573;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLPRTP 7
||||||
Db 91 RLPRTP 96

RESULT 15
G83465
conserved hypothetical protein PA1433 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: G83465
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83465
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-650 <STO>
A;Cross-references: UNIPROT:Q9I3R4; UNIPARC:UPI00000C537B; GB:AE004573; GB:AE004091; NI
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1433

Query Match 86.8%; Score 33; DB 2; Length 650;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLPRTP 7
||||||
Db 197 RLPRTP 202

Search completed: April 4, 2006, 13:17:20
Job time : 3.14529 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: April 4, 2006, 13:01:37 ; Search time 5.35079 Seconds
(without alignments)
922.986 Million cell updates/sec

Title: US-10-632-388-294
Perfect score: 38
Sequence: 1 RRLPRTP 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	100.0	96	2	Q4TGE7_TETNG	Q4tge7 tetraodon n
2	38	100.0	289	2	Q79FW0_MYCTU	Q79fw0 mycobacteri
3	38	100.0	289	2	Q7U179_MYCBO	Q7u179 mycobacteri
4	38	100.0	295	2	Q8VKD6_MYCTU	Q8vkd6 mycobacteri
5	38	100.0	736	2	Q4SKN8_TETNG	Q4skn8 tetraodon n
6	38	100.0	758	2	Q5B9I5_EMENI	Q5b9i5 aspergillus
7	38	100.0	819	2	Q4RWT8_TETNG	Q4rwt8 tetraodon n
8	38	100.0	1171	2	Q5B9Q6_EMENI	Q5b9q6 aspergillus
9	38	100.0	1337	2	Q6PCS4_BRARE	Q6pcs4 brachydanio
10	38	100.0	1538	2	Q5B9W9_EMENI	Q5b9w9 aspergillus
11	38	100.0	1581	2	Q5AST2_EMENI	Q5ast2 aspergillus
12	38	100.0	1590	2	O30915_STRFL	O30915 streptomyce
13	38	100.0	5830	2	Q50E74_STRFL	Q50e74 streptomyce
14	36	94.7	485	1	ZDHC1_HUMAN	Q8wtx9 homo sapien
15	36	94.7	678	1	PPN1_CRYNE	Q5kh67 cryptococcu
16	36	94.7	678	2	Q55ST2_CRYNE	Q55st2 cryptococcu
17	35	92.1	500	2	Q93HP6_STRAW	Q93hp6 streptomyce
18	35	92.1	678	2	Q92NV8_RHIME	Q92nv8 rhizobium m
19	35	92.1	701	2	Q61ZK0_CAEBR	Q61zk0 caenorhabdi
20	35	92.1	761	2	Q22005_CAEEL	Q22005 caenorhabdi
21	35	92.1	763	2	Q8I113_CAEEL	Q8i113 caenorhabdi
22	35	92.1	774	2	Q8BLR8_MOUSE	Q8blr8 mus musculu
23	35	92.1	780	2	Q8BLJ2_MOUSE	Q8blj2 mus musculu
24	35	92.1	797	2	Q86R99_DROME	Q86r99 drosophila
25	35	92.1	927	2	Q86RA0_DROME	Q86ra0 drosophila
26	35	92.1	1101	2	Q8MV37_DROME	Q8mv37 drosophila
27	35	92.1	1102	2	Q9VUX6_DROME	Q9vux6 drosophila
28	35	92.1	1144	2	Q8MT32_DROME	Q8mt32 drosophila
29	35	92.1	1145	2	Q8IQN4_DROME	Q8iqn4 drosophila
30	35	92.1	1182	2	Q5PQJ5_RAT	Q5pqj5 rattus norv
31	35	92.1	1252	2	Q59H86_HUMAN	Q59h86 homo sapien
32	35	92.1	1261	2	O15463_HUMAN	O15463 homo sapien
33	35	92.1	1261	2	Q52LW3_HUMAN	Q52lw3 homo sapien
34	35	92.1	1261	2	Q5VYZ0_HUMAN	Q5vyz0 homo sapien
35	35	92.1	1266	2	Q8CGF1_MOUSE	Q8cgf1 mus musculu
36	35	92.1	1335	1	RRPO_FXMV	P22168 foxtail mos
37	35	92.1	1335	2	Q8BB03_FXMV	Q8bb03 foxtail mos
38	34	89.5	154	2	Q8LLN1_ORYSA	Q8lln1 oryza sativ
39	34	89.5	160	2	Q9L1M3_STRCO	Q9l1m3 streptomyce
40	34	89.5	188	2	Q7Y016_ORYSA	Q7y016 oryza sativ
41	34	89.5	251	2	Q7VUY9_BORPE	Q7vuy9 bordetella
42	34	89.5	251	2	Q7WBD1_BORPA	Q7wbd1 bordetella
43	34	89.5	251	2	Q7WMV0_BORBR	Q7wmv0 bordetella
44	34	89.5	334	2	Q8DL84_SYNEL	Q8dl84 synechococc
45	34	89.5	411	2	Q9HND5_HALSA	Q9hnd5 halobacteri

ALIGNMENTS

RESULT 1
Q4TGE7_TETNG
ID Q4TGE7_TETNG PRELIMINARY; PRT; 96 AA.
AC Q4TGE7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF3748, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00001141001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetradontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01003748; CAF88035.1; -; Genomic_DNA.
FT NON_TER 1 1
SQ SEQUENCE 96 AA; 10508 MW; 9CC68444221CA322 CRC64;

Query Match 100.0%; Score 38; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLPRTP 7
|||||||
Db 38 RRLPRTP 44

RESULT 2

```
Q79FW0_MYCTU
ID   Q79FW0_MYCTU PRELIMINARY;      PRT;   289 AA.
AC   Q79FW0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   PROBABLE AMINO ACID AMINOTRANSFERASE (EC 2.6.1.-).
GN   Name=pabC; OrderedLocusNames=Rv0812;
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex.
OX   NCBI_TaxID=1773;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA   Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA   Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA   Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA   Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA   Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
DR   EMBL; BX842574; CAE55325.1; -; Genomic_DNA.
DR   TubercuList; Rv0812; -.
DR   GO; GO:0008483; F:transaminase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR001544; Aminotrans_IV.
DR   Pfam; PF01063; Aminotran_4; 1.
DR   ProDom; PD001961; Aminotrans_IV; 1.
KW   Aminotransferase; Complete proteome; Transferase.
SQ   SEQUENCE   289 AA;  31122 MW;  A384C2289727FC80 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        266 RRLPRTP 272


RESULT 3
Q7U179_MYCBO
ID   Q7U179_MYCBO PRELIMINARY;      PRT;   289 AA.
AC   Q7U179;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   PROBABLE AMINO ACID AMINOTRANSFERASE (EC 2.6.1.-).
GN   OrderedLocusNames=Mb0835;
OS   Mycobacterium bovis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex.
OX   NCBI_TaxID=1765;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AF2122/97;
RX   MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA   Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA   Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA   Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA   Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT   "The complete genome sequence of Mycobacterium bovis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR   EMBL; BX248336; CAD93697.1; -; Genomic_DNA.
DR   GO; GO:0008483; F:transaminase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR001544; Aminotrans_IV.
DR   Pfam; PF01063; Aminotran_4; 1.
DR   ProDom; PD001961; Aminotrans_IV; 1.
KW   Aminotransferase; Complete proteome; Transferase.
SQ   SEQUENCE   289 AA;  31140 MW;  14A941455727FBED CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        266 RRLPRTP 272


RESULT 4
Q8VKD6_MYCTU
ID   Q8VKD6_MYCTU PRELIMINARY;      PRT;   295 AA.
AC   Q8VKD6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Aminotransferase, class IV.
GN   Name=dat; OrderedLocusNames=MT0833;
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex.
OX   NCBI_TaxID=1773;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CDC 1551 / Oshkosh;
RX   MEDLINE=22206494; PubMed=12218036;
RX   DOI=10.1128/JB.184.19.5479-5490.2002;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA   Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA   Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA   Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA   Fraser C.M.;
RT   "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   J. Bacteriol. 184:5479-5490(2002).
DR   EMBL; AE000516; AAK45075.1; -; Genomic_DNA.
DR   TIGR; MT0833; -.
DR   GO; GO:0008483; F:transaminase activity; IEA.
DR   GO; GO:0016740; F:transferase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR001544; Aminotrans_IV.
DR   Pfam; PF01063; Aminotran_4; 1.
DR   ProDom; PD001961; Aminotrans_IV; 1.
KW   Aminotransferase; Transferase.
SQ   SEQUENCE   295 AA;  31795 MW;  591421FEEE16F04E CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 295;
  Best Local Similarity   100.0%;  Pred. No. 39;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        272 RRLPRTP 278


RESULT 5
Q4SKN8_TETNG
ID   Q4SKN8_TETNG PRELIMINARY;      PRT;   736 AA.
AC   Q4SKN8;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
```

```
DE   Chromosome undetermined SCAF14565, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00016637001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014565; CAF98794.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER     736    736
SQ   SEQUENCE   736 AA;  82661 MW;  650BAE00C47918C9 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 736;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        406 RRLPRTP 412


RESULT 6
Q5B9I5_EMENI
ID   Q5B9I5_EMENI PRELIMINARY;      PRT;   758 AA.
AC   Q5B9I5;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=AN2795.2;
OS   Aspergillus nidulans FGSC A4.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=227321;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FGSC A4;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Genome Sequence of Aspergillus nidulans.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACD01000049; EAA63229.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   758 AA;  83393 MW;  A474F8AFB4CCE12B CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 758;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        663 RRLPRTP 669


RESULT 7
Q4RWT8_TETNG
ID   Q4RWT8_TETNG PRELIMINARY;      PRT;   819 AA.
AC   Q4RWT8;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 15 SCAF14981, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00027697001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014981; CAG07144.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER     819    819
```

```
SQ   SEQUENCE   819 AA;  91287 MW;  8DF4D05210E5A623 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 819;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        483 RRLPRTP 489


RESULT 8
Q5B9Q6_EMENI
ID   Q5B9Q6_EMENI PRELIMINARY;      PRT;  1171 AA.
AC   Q5B9Q6;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=AN2724.2;
OS   Aspergillus nidulans FGSC A4.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=227321;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FGSC A4;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Genome Sequence of Aspergillus nidulans.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACD01000048; EAA63022.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1171 AA;  127986 MW;  C72CBEAAFD7DEB56 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 1171;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        531 RRLPRTP 537


RESULT 9
Q6PCS4_BRARE
ID   Q6PCS4_BRARE PRELIMINARY;      PRT;  1337 AA.
AC   Q6PCS4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein zgc:63950.
GN   ORFNames=zgc:63950;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC059184; AAH59184.1; -; mRNA.
DR   ZFIN; ZDB-GENE-031010-44; zgc:63950.
DR   GO; GO:0019992; F:diacylglycerol binding; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0005096; F:GTPase activator activity; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   InterPro; IPR002219; DAG_PE-bind.
DR   InterPro; IPR000198; RhoGAP.
DR   InterPro; IPR012294; TFIID_C/glycos_N.
DR   Pfam; PF00130; C1_1; 1.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00109; C1; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR   PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR   PROSITE; PS50238; RHOGAP; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1337 AA;  147665 MW;  2061AE2507489079 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 1337;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        623 RRLPRTP 629


RESULT 10
Q5B9W9_EMENI
ID   Q5B9W9_EMENI PRELIMINARY;      PRT;  1538 AA.
AC   Q5B9W9;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
```

```
DE   Hypothetical protein.
GN   ORFNames=AN2661.2;
OS   Aspergillus nidulans FGSC A4.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=227321;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FGSC A4;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Genome Sequence of Aspergillus nidulans.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACD01000046; EAA63063.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1538 AA;  170050 MW;  7252FCEEFDA6EDA1 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 1538;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        687 RRLPRTP 693


RESULT 11
Q5AST2_EMENI
ID   Q5AST2_EMENI PRELIMINARY;      PRT;  1581 AA.
AC   Q5AST2;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=AN8648.2;
OS   Aspergillus nidulans FGSC A4.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=227321;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=FGSC A4;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Genome Sequence of Aspergillus nidulans.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACD01000158; EAA60682.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1581 AA;  174385 MW;  E793EAAA2783CC37 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 1581;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        662 RRLPRTP 668


RESULT 12
O30915_STRFL
ID   O30915_STRFL PRELIMINARY;      PRT;  1590 AA.
AC   O30915;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Daptomycin biosynthetic protein subunit (Fragment).
OS   Streptomyces filamentosus (Streptomyces roseosporus).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=67294;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=A21978.6;
RX   MEDLINE=98083067; PubMed=9422604;
RA   Mchenney M.A., Hosted T.J., Dehoff B.S., Rosteck P.R. Jr., Baltz R.H.;
RT   "Molecular cloning and physical mapping of the daptomycin gene cluster
RT   from Streptomyces roseosporus.";
RL   J. Bacteriol. 180:143-151(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=A21978.6;
RA   McHenney M.A., DeHoff B.S., Rosteck P.R. Jr., Baltz R.H.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC       family.
DR   EMBL; AF021263; AAB96629.1; -; Genomic_DNA.
DR   HSSP; P14687; 1AMU.
DR   GO; GO:0048037; F:cofactor binding; IEA.
DR   GO; GO:0016874; F:ligase activity; IEA.
DR   GO; GO:0031177; F:phosphopantetheine binding; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   InterPro; IPR010071; AA_adenyl_dom.
DR   InterPro; IPR009081; ACP_like.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; Condensatn.
DR   InterPro; IPR006163; Phsppanteth_bind.
DR   InterPro; IPR006162; Ppantne_S.
DR   Pfam; PF00501; AMP-binding; 2.
DR   Pfam; PF00668; Condensation; 2.
```

```
DR   Pfam; PF00550; PP-binding; 1.
DR   PRINTS; PR00154; AMPBINDING.
DR   TIGRFAMs; TIGR01733; AA-adenyl-dom; 1.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
DR   PROSITE; PS00455; AMP_BINDING; 2.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
KW   Phosphopantetheine.
FT   NON_TER       1      1
FT   NON_TER    1590   1590
SQ   SEQUENCE   1590 AA;  171015 MW;  33C5CB3D26983E72 CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 1590;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db        871 RRLPRTP 877


RESULT 13
Q50E74_STRFL
ID   Q50E74_STRFL PRELIMINARY;      PRT;  5830 AA.
AC   Q50E74;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Peptide synthetase 1.
GN   Name=dptA;
OS   Streptomyces filamentosus (Streptomyces roseosporus).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=67294;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=NRRL 11379;
RA   Miao V., Coeffet-LeGal M.-F., Brian P., Brost R., Penn J., Whiting A.,
RA   Martin S., Ford R., Parr I., Bouchard M., Silva C.J., Wrigley S.K.,
RA   Baltz R.H.;
RT   "Daptomycin biosynthesis in Streptomyces roseosporus: cloning and
RT   analysis of the gene cluster and revision of peptide
RT   stereochemistry.";
RL   Microbiology 151:1507-1523(2005).
CC   -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC       family.
DR   EMBL; AY787762; AAX31557.1; -; Genomic_DNA.
DR   InterPro; IPR010071; AA_adenyl_dom.
DR   InterPro; IPR009081; ACP_like.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; Condensatn.
DR   InterPro; IPR010060; NRPS_synth.
DR   InterPro; IPR000169; Pept_cys_AS.
DR   InterPro; IPR006163; Phsphopanteth_bd.
DR   InterPro; IPR006162; Ppantne_S.
DR   Pfam; PF00501; AMP-binding; 5.
DR   Pfam; PF00668; Condensation; 6.
DR   Pfam; PF00550; PP-binding; 5.
DR   PRINTS; PR00154; AMPBINDING.
DR   TIGRFAMs; TIGR01733; AA-adenyl-dom; 5.
DR   TIGRFAMs; TIGR01720; NRPS-para261; 1.
DR   PROSITE; PS50075; ACP_DOMAIN; 5.
DR   PROSITE; PS00455; AMP_BINDING; 5.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; 5.
DR   PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW   Phosphopantetheine.
SQ   SEQUENCE   5830 AA;  624177 MW;  B53232641B6EA34C CRC64;

  Query Match             100.0%;  Score 38;  DB 2;  Length 5830;
  Best Local Similarity   100.0%;  Pred. No. 1e+03;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              |||||||
Db       4655 RRLPRTP 4661


RESULT 14
ZDHC1_HUMAN
ID   ZDHC1_HUMAN    STANDARD;      PRT;   485 AA.
AC   Q8WTX9; O15461;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Zinc finger DHHC domain containing protein 1 (Zinc finger protein 377)
DE   (DHHC-domain-containing cysteine-rich protein 1).
GN   Name=ZDHHC1; Synonyms=C16orf1, ZNF377;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Uterus;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE OF 1-293, AND TISSUE SPECIFICITY.
RC   TISSUE=Pancreas;
RX   MEDLINE=99321009; PubMed=10395086; DOI=10.1023/A:1006932522197;
RA   Putilina T., Wong P., Gentleman S.;
RT   "The DHHC domain: a new highly conserved cysteine-rich motif.";
RL   Mol. Cell. Biochem. 195:219-226(1999).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- TISSUE SPECIFICITY: Expressed at high levels in fetal lung, kidney
CC       and heart. Expressed at lower levels in adult pancreas and lung.
CC   -!- SIMILARITY: Contains 1 DHHC-type zinc finger.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; BC021908; AAH21908.1; -; mRNA.
DR   EMBL; U90653; AAB86591.2; -; mRNA.
DR   HSSP; P28814; 1BW4.
DR   Ensembl; ENSG00000159714; Homo sapiens.
DR   HGNC; HGNC:17916; ZDHHC1.
DR   GO; GO:0003677; F:DNA binding; NAS.
DR   GO; GO:0005515; F:protein binding; NAS.
DR   InterPro; IPR001594; Znf_DHHC.
DR   Pfam; PF01529; zf-DHHC; 1.
DR   ProDom; PD003041; Znf_DHHC; 1.
DR   PROSITE; PS50216; ZF_DHHC; 1.
KW   Metal-binding; Transmembrane; Zinc; Zinc-finger.
FT   TRANSMEM     53     73       Potential.
```

```
FT   TRANSMEM     78     98       Potential.
FT   TRANSMEM    186    206       Potential.
FT   TRANSMEM    242    262       Potential.
FT   ZN_FING     134    184       DHHC-type.
SQ   SEQUENCE   485 AA;  54818 MW;  6B75B077D782F358 CRC64;

  Query Match             94.7%;  Score 36;  DB 1;  Length 485;
  Best Local Similarity   85.7%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              ||:||||
Db        422 RRIPRTP 428


RESULT 15
PPN1_CRYNE
ID   PPN1_CRYNE     STANDARD;      PRT;   678 AA.
AC   Q5KH67;
DT   10-MAY-2005 (Rel. 47, Created)
DT   10-MAY-2005 (Rel. 47, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Endopolyphosphatase (EC 3.6.1.10).
GN   Name=PPN1; OrderedLocusNames=CNE01080;
OS   Cryptococcus neoformans (Filobasidiella neoformans).
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC   Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX   NCBI_TaxID=5207;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=JEC21;
RX   PubMed=15653466; DOI=10.1126/science.1103773;
RA   Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA   Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA   Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA   D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA   Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA   Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA   Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
RA   Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA   Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
RA   Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA   Fraser C.M., Hyman R.W.;
RT   "The genome of the basidiomycetous yeast and human pathogen
RT   Cryptococcus neoformans.";
RL   Science 307:1321-1324(2005).
CC   -!- FUNCTION: Catalyzes the hydrolysis of inorganic polyphosphate
CC       (poly P) chains of many hundreds of phosphate residues into
CC       shorter lengths (By similarity).
CC   -!- CATALYTIC ACTIVITY: Polyphosphate + n H(2)O = (n+1)
CC       oligophosphate.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Vaculoar.
CC   -!- PTM: Processing by proteases in the vacuole may be required for
CC       activation (By similarity).
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; AE017345; AAW43547.1; -; Genomic_DNA.
DR   InterPro; IPR004843; M-pesterase.
DR   Pfam; PF00149; Metallophos; 1.
KW   Complete proteome; Glycoprotein; Hydrolase; Signal-anchor;
KW   Transmembrane; Vacuole.
FT   TOPO_DOM      1      2       Cytoplasmic (Potential).
FT   TRANSMEM      3     23       Signal-anchor for type II membrane
FT                                protein (Potential).
FT   TOPO_DOM     24    678       Vacuolar (Potential).
FT   CARBOHYD    138    138       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    369    369       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    447    447       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    591    591       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    616    616       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   678 AA;  77209 MW;  8B3C696309AC617D CRC64;

  Query Match             94.7%;  Score 36;  DB 1;  Length 678;
  Best Local Similarity   85.7%;  Pred. No. 2.4e+02;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLPRTP 7
              ||:||||
Db        168 RRIPRTP 174


Search completed: April  4, 2006, 13:15:09
Job time : 8.35079 secs
```